CHILD DEVELOPMENT

Vol. 25
December, 1954
No. 4

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS

OF THE

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Contents for December

Information for Contributors and Subscribers

Child Development, one of three publications issued by the Society for Research in Child Development, Inc., is issued quarterly in March, June, September and December. The subscription price is $6.00 per year, postpaid.

Child Development Abstracts and Bibliography is now in its twenty-seventh volume. This publication is issued February through December—six numbers to a volume. The subscription price is $5.00 per year, postpaid.

Monographs of the Society for Research in Child Development, Inc., issued irregularly during the year, consists primarily of detailed critical reviews and seriatim studies. The subscription price is $4.00 per yearly volume.

The three publications of the Society are available at the special price of $12.00 per year.

A limited number of back issues of all publications is available at reduced prices. A mimeographed list may be had upon request.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (Supplement to the *Psychological Bulletin,* 1952, 49, No. 4, Part 2) and *A Manual of Style* (University of Chicago Press, 1950) in the preparation of manuscripts.

Manuscripts and *all business correspondence* relating to the publications should be addressed to Child Development Publications, 1341 Euclid, University of Illinois, Champaign, Illinois.

CHILD DEVELOPMENT

Vol. 25

December, 1954

No. 4

WILLIAM E. MARTIN, *Editor*
University of Illinois

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
1341 Euclid, University of Illinois
Champaign, Illinois

Made in United States of America

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

THE ROLE OF MEMBER IN CLUBS OF LOWER-CLASS AND MIDDLE-CLASS ADOLESCENTS[1]

HENRY S. MAAS
University of California, Berkeley

Is the type of role-adaptation adolescents make as members of formally organized groups related to their social class affiliation? Specifically, is the role of member fulfilled in different ways in neighborhood center teenage clubs in the slums and in middle-class areas? In the sizeable body of empirical studies on children's group behavior and on social class influences upon adolescent life, one finds few generalizations that bear directly upon this problem.

In their informal associations with peers, adolescents value certain behaviors and attributes that vary with social class membership (6, 14, 16). In addition to peer culture mores, however, role-adaptation in teen-age clubs is theoretically dependent upon the formal structure of such groups, including a president or chairman, members who are clearly so designated, and an adult leader whom the neighborhood center assigns to foster youth's practice in democratic decision-making and for other purposes. Slum area street gangs, which may be in their own way as clearly structured as the neighborhood center clubs, have been vividly described by Thrasher (17), Whyte (19), and others, but comparable accounts on middle-class neighborhood youth groups are lacking, and none of these groups, of course, has a responsible adult regularly in attendance. Membership in formally organized groups and other components in the social participation patterns of middle- and lower-class youth have been studied (12, 13), but research on role-adaptation *within* such groups has failed to utilize social class as an independent variable. Instead, when the social class concept has been used, it has been for purposes of experimental control, as for example in Lippitt's

[1] The writer is indebted to the following persons who were members of the year's graduate seminar in which this study was executed: Robert T. Blazejack, Evelio Grillo, Robert J. Hagest, Margie B. Herman, Genevieve C. Hoffman, Donald Z. Miller, Charles O'Shea, Irving M. Piliavin, Joseph H. Solis, Memya Thoren, and William B. Ward. The study was done as a Group Research Master's Project at the School of Social Welfare, University of California, Berkeley, under the writer's faculty supervision. The graduate students served as group observers and in many other roles in the course of the investigation.

first laboratory study on children's groups (8).[2] Lacking systematic empirical inquiry, the relationship of social class status to role-adaptation in formally organized groups is open to such speculative generalization as: "The worker [working class] group tends to remain relatively unstructured—roles are ill-defined and quite fluid, preventing fruitful specialization and growth (2, p. 29)."

Drawing upon Davis' research on social class theory and adolescent development (3, 4), one might postulate that lower-class teenage club members express more aggression, engage more frequently in digressions from the central activity of a meeting, and are generally less collaborative as a total group during club meetings than are comparable middle-class club members. Prior research, in Chicago (1, 10, 11, 15), of which the current study is a continuation, reveals that this simple formulation does not stand up under testing. Rather, the prior research suggests that the roles of member, club president, and adult leader are quite differently perceived by slum area and middle-class neighborhood center adolescents. A proposition on role inferred from narrative records of club meetings and interviews with a cross-section of members, early in the Chicago research, seemed to be confirmed when all the data on twelve clubs were in. Since the current study aims to re-test, with more refined research techniques, certain parts of this formulation, its quotation here is appropriate:

> In the middle-class groups, pre- and early adolescents seem primarily directed toward working out relationships with peers. At home their proximate relationships with parents in joint activities—going places and doing things together—and the freedom to express verbally ther negative reactions, to gainsay parental requests, and to manipulate parents in ways that suggest their equal status and power with parents make relations with the adult adviser in youth group settings developmentally less significant. In the lower-class groups, however, the adult adviser becomes a more significant and conflictful person, someone in a sense who is supra-human—to be tried, or challenged more or less openly. Relations among peers in the lower-class groups are characterized by repeated obvious attempts at status seeking. . . . This pattern in the lower-class groups may be traced to the increasing distance lower-class boys and girls in growing up have attained from adults; earlier on their own (as Davis and others have shown), they seem to have gotten out of communication with adults. Physical aggression in their

[2] Lippitt describes the children in both groups as though none came from lower-class families. "With one or two exceptions, equally true for both clubs, we may roughly generalize . . . [for all subjects] : (1) a home of good socio-economic status, few siblings, and consistently moderate discipline; (2) an unorganized neighborhood play life with no hint of gang membership; (3) rather inactive membership in a Club pack; and (4) membership in a progressively inclined university school system" (8, p. 65). With subjects of this type of background, one may question Lippitt's generalizing, albeit with caution, about behavior of "the authoritarian group members" and "the democratic group" (8, pp. 190-191). What might have happened had some lower-class children been included in each of these groups?

> homes, when parents are at home together with their pre- and early adolescent children, continues to assure them the adult is boss. Personal relationships are seen in a hierarchical framework, definitely non-equalitarian and non-collaborative, and the power struggle between adults and early adolescents carries over into the youth groups, with much hostility, overt, covert, and smothered (11, p. 12).

For the current study, the concept of role has been operationalized as having structural and functional properties. Structurally, the role of member is fulfilled in reciprocal relationships with the roles of adult leader, president, and other member. Functionally, each role may include behaviors that are primarily collaborative, in the process of democratic decision-making; hostilely aggressive, in the authoritarian tradition of might-is-right; or digressive, an asocial or anarchic mode of interaction. The psychological counterparts of these three modes of social process appear in Horney's statement that "a child can move *toward* people, *against* them or *away from* them (7, p. 42)." In a teenage club, these three modes of interaction can be directed by a member in any one of three structural directions: to the adult leader, president, or other member(s). The functional and structural properties of role seem adequately operationalized if every interaction is seen as having conjointly a mode and a direction.

If the role of member and its reciprocal roles are differentially perceived by lower-class and middle-class adolescents, the modes and directions of interaction that members initiate and receive should reflect this differential pattern. Members may sanction or taboo more or less interaction with the adult leader or the president. The mode of such interaction theoretically depends upon the perception of each role. "Toward" behavior or collaboration is assumed to imply respect or trust; as a group-integrative or centripetal mode of interaction that leads to the resolution of an issue, it includes the behaviors of suggesting, asking, yielding, approving, deciding, and the incorporating of oppositional or excluded persons. (See Figure 1.) "Against" behavior or hostile aggression implies distrust or threat and is manifest by compelling or forcing, resisting, and either verbal or non-verbal attack; interaction which aims to superordinate the initiator and subordinate the object of the interaction, such as compelling, is thus considered hostilely aggressive. "Away from" behavior or digression is group-disintegrative or centrifugal interaction in the process of movement away from threatening or disapproved others and/or central activity of the group at a given time; it may appear as subgrouping or fragmentation of members who engage in tangential "horse-play" or other non-group relevant interaction during discussion on the planning of future events. Collaboration, aggression, and digression are three modes of behavior that may be incorporated with greater or lesser frequency into the role of member in different groups and directed with greater or lesser frequency to the correlative roles of adult leader, club president, or other member.

Specifically, then, drawing upon what is known and postulated about lower-class and middle-class adolescent group life, the following hypotheses are proposed on the lower-class, as compared with the middle-class, club member:

Hypothesis 1: In lower-class groups, the member directs to the adult leader (a) more collaboration, (b) more aggression, and (c) more digression. The adult is developmentally "a more significant and conflictful person" for him (11, p. 12).

Hypothesis 2: In lower-class groups, the member directs to the president (a) less collaboration, (b) less aggression, and (c) less digression. More obvious in his status-seeking among peers (11, p. 12), the lower-class club member is likely either quickly to conform with or to ignore a president, depending on his true status in the peer society; whether high or low, he is not likely to be the focus of more frequent interactions than his counterpart in middle-class groups who may be either approved or disapproved, but in either case is seen as the occupant of a position that must be dealt with, may be widely aspired to, and seems essential to the achievement of club and neighborhood center goals. Such postulates are at least in line with Davis' concept of "social anxiety" among middle-class youth (3).

Hypothesis 3: In lower-class groups, the member directs to other members (a) less collaboration, (b) more aggression, and (c) more digression. This follows from the characteristically more overt expression of aggression and lesser degree of impulse control found among lower-class youth today (3).

Each of these hypotheses is the converse of a hypothesis on the role of member in middle-class adolescent clubs.

Methods

To test the hypotheses, ten clubs averaging a total attendance of 126 adolescent members were selected and paired in the San Francisco Bay Area. An observational instrument (Figure 1) was developed and used in three successive mid-year meetings in each club.

For the selection of clubs, criteria which provide control of factors theoretically related to members' role-adaptation were set up, members' social class status being the independent variable. The criteria that could be met in the field are as follows: (a) all clubs have scheduled weekly meetings in a neighborhood center attracting primarily either slum area or middle-class adolescents; (b) all clubs have an agency-designated adult leader regularly in attendance; (c) all club meetings are presided over by an adolescent in the role of president; (d) all club meetings include a forum-type discussion for the members' planning of their own programs, thus excluding from this study all groups with highly ritualized procedures and agency-imposed programs as in some Scout troops; (e) all clubs are organ-

Observer _____ Club Name _____

Time Started _____ Date _____

Time Ended _____ No. Members _____

Observation No. 1 2 3 (circle) Co-observer _____

	Leader		President		Member(s)			
TOWARD (collaboration)	P	M	L	M	L	P	M	
1. suggest								
2. ask for								
3. yield								
4. approve								
5. decide								
6. incorporate								
AGAINST (aggression)	P	M	L	M	L	P	M	
1. compel, force								
2. resist								
3. attack ..... verbal								
non-verbal								
AWAY FROM (digression)	P	M	L	M	L	P	M	
1. non-group relevant interaction verbal								
non-verbal								
2. silence								

FIGURE 1—Observational tally-sheet used by observers for each six-minute observation period.

ized and meeting for a minimum of four months by the time that research observations begin, thus excluding from this study all very newly organized clubs still in the process of initial group formation and role-definition; and (f_1) all clubs are composed of either boys or girls, ages 12 to 16, with a median attendance of not fewer than 6 and not more than about 20 mem-

bers, but (f_2) all clubs paired for co-observation are to be of the same sex, not more than about a year apart in median age of members, not too disparate in numbers attending, and each composed predominantly of lower-class or middle-class adolescents. These criteria are in certain ways a broadening of the initial theoretically-set controls, but limits had to be expanded in view of the clubs available for pairing in the 21 canvassed urban neighborhood centers. Matching of the adult leaders in paired clubs was considered a theoretically questionable and operationally infeasible procedure in this field investigation, if only because of evidence that the role of adult leader is more a function of the structure and process of the group than a direct correlate of any as yet isolated personal or ego-structure factor in the leader as an individual (9). Data on the leaders of the ten clubs, obtained after their selection, reveal that the adults working with the lower-class and the middle-class adolescents were roughly comparable as to prior experience as adult leader, educational level, and prior personal experience as member of a youth group.

Lower-class and middle-class clubs were paired, as shown in Table 1. Each pair of clubs was observed by the same team of two observers, thus reducing the possibility of ascribing observed differences to differences among observers. Both observers in each team were of the same sex as the members of the paired clubs they observed.

TABLE 1

AGE, SEX, ATTENDANCE, PARENTAL OCCUPATION, AND SOCIAL CLASS OF ADOLESCENT MEMBERS IIN TEN PAIRED CLUBS

Club	*Age Range* (years)	*Median Age*	*Sex*	*Median Attendance* ($N=126$)	*Median Parental Occupation**	*Social Class*
A	12–13	12.8	M	10	2	lower
B	11–13	12.5	M	6	6	middle
C	14–16	14.7	M	15	3	lower
D	12–14	13.5	M	18	6	middle
E	13–16	14.0	M	8	3	lower
F	13–14	13.5	M	21	7	middle
G	16–17	16.5	M	8	4	lower
H	14–17	15.7	M	12	7	middle
I	13–15	13.8	F	17	4	lower
J	12–13	12.6	F	11	7	middle
Lower class (mean) ...		14.4	—	11.6	3.2	—
Middle class (mean) ..		13.6	—	13.6	6.6	—

* See footnote 3 for rating scale.

Social class estimates of club members were made first by comparative evaluation of neighborhood (dwelling area) and then by occupational rating[3] of club members' parents, the latter rating for all clubs confirming the former estimates. While all the middle-class clubs were predominantly Caucasian and non-ethnic, clubs A, C, and I (see Table 1) were composed of Negroes, and clubs E and G met in a neighborhood center that served a large percentage of Italian ethnics.[4]

TABLE 2

RELIABILITY OF TALLY-SHEET OBSERVERS INDICATED BY RANK ORDER CORRELATIONS OF TYPE OF MODE AND DIRECTION OF INTERACTIONS

Observers	Pre-test 1		Pre-test 2		Pre-test 3		Mean
	Type of mode	*Direction*	*Type of mode*	*Direction*	*Type of mode*	*Direction*	
I – II	—*	.69	.84	.71	.68	.71	.73
I – IV	—*	.90	.64	.76	.74	.84	.78
II – IV	.74	.61	.37	.38	.82	.50	.57
III – V	.97	.88	.78	.88	.63	.90	.84
VI – VII	.94	.45	.92	.83	.70	.74	.76
VIII – IX	.97	.79	.73	.77	.86	.73	.81

* Coefficient not obtained because of error by Observer I in recording.

In each team of co-observers, one wrote a narrative record and one used the observational tally-sheet (Figure 1) to obtain three spaced six-minute observations in each of three successive meetings. (Data from the narrative records are not discussed in this report.) To obtain samples of early, middle, and later interactions in each meeting of each club, the first six-minute observation on a tally-sheet began five minutes after each meeting had begun; the second, twenty minutes after the beginning of each meeting; and the third, as close to the termination of each meeting as possible. In the development of the tally-sheet, its validity was judged non-quantitatively in pre-tests on club meetings role-played by members of the research group; pre-planned interaction patterns and narrative reports of these "clubs" in

[3] Ratings are based on a scale developed for the U.S. Bureau of the Census (20). Higher ratings are for proprietors, managers, and professionals, 8 or 7; sales and clerical workers, 6; and skilled craftsmen and foremen, 5. Lower ratings are for semi-skilled and unskilled workers: operatives, 4; household and service workers, 3; laborers, 2; and relief recipients, 1.

[4] Minority group ethnicity is a factor in urban lower-class culture that Warner has built into his social class theory (18, pp. 118-140) and that Davis and Havighurst have found for Negroes and Caucasians in no way obfuscates social class differences in child-rearing (5).

action were compared with the tally-sheet records. The reliability of an instrument like the tally-sheet is determined primarily by the reliability of its users. Rank order correlation of the thirteen types of interaction (horizontal items, Figure 1) and the seven direction totals (vertical items, Figure 1) revealed that all but two of the pre-test observers (II and IV) could use the tally-sheets with some degree of objectivity (Table 2). Observers I, III, V, VII, and IX were chosen to do the tally-sheet recording in the field; the others, and observer X, became the narrative report co-observers.

TABLE 3

FREQUENCY BY MODE AND DIRECTION OF INTERACTIONS INITIATED BY MEMBERS IN LOWER-CLASS AND MIDDLE-CLASS CLUBS

Direction *Mode*	*Lower class*	*Middle class*	*Total*	*Chi-square value*
To Adult Leader:				
Collaboration	94	64	158	5.70*
Aggression	44	31	75	2.25
Digression	10	10	20	...
To Club President:				
Collaboration	50	85	135	9.07†
Aggression	18	42	60	9.60†
Digression	5	5	10	...
To Other Member(s):				
Collaboration	208	240	448	2.29
Aggression	84	55	139	6.05*
Digression	270	243	513	1.42
TOTAL	783	775	1558	36.38‡

Significant at: * .02 level † .01 level ‡ .001 level

FINDINGS AND CONCLUSIONS

The role of member is differently fulfilled and apparently differently perceived by the lower- and the middle-class adolescents, as may be seen in Tables 3 and 4.

Hypothesis 1: The lower-class adolescent club member directs more collaborative interactions to the adult leader than does the middle- class member (Table 3). The data do not support parts (b) and (c) of this hypothesis in regard to more aggression and digression directed to the adult leader in slum area groups. It should be noted that the adult leader reciprocally directs more collaborative interactions to lower-class than to middle-class members (Table 4).

Hypothesis 2: The middle-class adolescent club member directs more collaborative and more aggressive interactions to the president (Table 3). The data do not support part (c) of this hypothesis on digressive interactions. The president, in turn, is more collaborative and aggressive with members in middle-class than in lower-class groups (Table 4).

Hypothesis 3: The lower-class adolescent club member is more aggressive with other members (Table 3). The data do not support part (a) of this hypothesis on less collaboration among lower-class adolescents. There

TABLE 4

FREQUENCY BY MODE AND DIRECTION OF INTERACTIONS RECEIVED BY MEMBERS IN LOWER-CLASS AND MIDDLE-CLASS CLUBS

Direction — *Mode*	*Lower class*	*Middle class*	*Total*	*Chi-square value*
From Adult Leader:				
Collaboration	249	164	413	21.40‡
Aggression	49	51	100	...
Digression	7	12	19	1.12
From Club President:				
Collaboration	135	221	356	17.10‡
Aggression	48	112	160	23.10‡
Digression	10	20	30	2.94**
From Other Member(s):				
Collaboration	208	240	448	0.57
Aggression	84	55	139	7.30†
Digression	270	243	513	2.70**
TOTAL	1060	1118	2178	76.23‡

Significant at: **.10 level †.01 level ‡.001 level

is some slight evidence (Table 4, at the .10 level of confidence) that there is more member-member digression in the lower-class clubs.

The data validate only parts of the initial proposition. The over-all pattern of quantitative findings is largely, however, in the expected directions. The initial formulation may thus be refined in the following ways:

1. The role of member among lower-class adolescents in neighborhood center clubs is fulfilled in a relationship to the adult leader that is comparable to the relationship the middle-class adolescent member has with his peer in the role of club president. In slum area clubs, the adult leader would seem to be a more significant figure to members than in middle-class neighborhood centers where adolescent members ostensibly have less need for interaction with the adult.

2. The role of member among middle-class adolescents apparently sanctions less aggression and digression in member-to-member interactions than does the comparable role among lower-class adolescents. This much squares with existing social class theory. It should be noted, however, that the total amount of aggression expressed by members in the two types of group is not significantly different, since the club president in the middle-class groups is the target (and initiator) of significantly more aggressive interactions than is his counterpart in the lower-class groups. One must conclude, therefore, that the role of member among both lower- and middle-class adolescents offers equal opportunity but different directional outlets for the expression of aggression in teenage clubs.

3. Differences found between lower-class and middle-class adolescents in the role of member and its relationships with the reciprocal roles of adult leader and club president may be explained in terms of differential cultural training and experience relevant to role-adaptation in formally organized groups. The initial proposition on differences in adolescent-adult and peer relationships in lower-class and middle-class child life is not invalidated by the findings of this study.

References

1. Bancker, J. Digressive behavior in middle- and lower-class organized youth groups. Unpublished master's thesis, Univ. Chicago, 1951.
2. Bellin, S., and Riessman, F., Jr. Education, culture, and the anarchic worker. *J. soc. Issues,* 1949, 5, 24-32.
3. Davis, A. Socialization and adolescent personality. *Yearb. nat. Soc. Stud. Educ.,* 1944, 43 (I), 198-216.
4. Davis, A., and Dollard, J. *Children of bondage.* Washington: American Council on Education, 1940.
5. Davis, A., and Havighurst, R. J. Social class and color differences in child-rearing. *Amer. sociol. Rev.,* 1946, 11, 698-710.
6. Hollingshead, A. B. *Elmtown's youth.* New York: Wiley, 1949.
7. Horney, K. *Our inner conflicts.* New York: Norton, 1945.
8. Lippitt, R. An experimental study of the effects of democratic and authoritarian group atmospheres. *Univ. Iowa Stud. Child Welf.,* 1940, 16, 43-195.
9. Maas, H. S. Personal and group factors in leaders' social perception. *J. abnorm. soc. Psychol.,* 1950, 45, 56-63.
10. Maas, H. S. Some social class differences in the family systems and group relations of pre- and early adolescents. *Child Develpm.,* 1951, 22, 145-152.

11. Maas, H. S. A study of group life in human development. *Hum. Develpm. Bull.*, 1951, 4, 11-15.

12. Macdonald, M., McGuire, C., and Havighurst, R. J. Leisure activities and the socioeconomic status of children. *Amer. J. Sociol.*, 1949, 54, 505-519.

13. McEntire, D. Leisure activities of youth in Berkeley, California. Berkeley Council of Social Welfare and School of Social Welfare, Univ. Calif., 1952.

14. Neugarten, B. L. Social class and friendship among school children. *Amer. J. Sociol.*, 1946, 51, 305-313.

15. Olson, V. E. Hostile aggression in organized youth groups. Unpublished doctor's dissertation, Univ. Chicago, 1951.

16. Pope, B. Prestige values in contrasting socioeconomic groups of children. *Psychiatry*, 1953, 16, 381-385.

17. Thrasher, F. M. *The gang*. Chicago: Univ. Chicago Press, 1927.

18. Warner, W. L. *Structure of American life*. Edinburgh: Univ. Edinburgh Press, 1952.

19. Whyte, W. F. *Street corner society*. Chicago: Univ. Chicago Press, 1943.

20. *Alphabetical index of occupations and industries*. Washington: Bureau of the Census, 1950.

EFFECTS OF TRAINING ON THE SRA PRIMARY MENTAL ABILITIES (PRIMARY) AND THE WISC[1]

HAROLD D. HOLLOWAY
Iowa Child Welfare Research Station[2]

There is at the present time a growth in number of various training programs for kindergarten children. These usually involve the use of some type of paper-and-pencil work-book. The programs are primarily designed "better to prepare" the child for his forthcoming academic experiences such as reading and arithmetic or to increase his intelligence, presumably as operationally defined by his responses to certain test stimuli and the conditions under which these responses are evoked. The present study is an attempt to appraise critically the effectiveness of one of these training programs.

The particular training program selected for study was one constructed by Thelma Gwinn Thurstone entitled *Learning to Think Series—The Red Book* (2). Certain statements in the teacher's manual (1) for the training program prompted the following four general non-statistical hypotheses. It was the purpose of this study to attempt to test these hypotheses using the SRA Primary Mental Abilities-Primary (3) (PMA) and the WISC (5) as separate criterion scores.

---

[1] This article represents part of a Ph.D. dissertation completed in 1953 at the Iowa Child Welfare Research Station, State University of Iowa. Acknowledgment is expressed to the late Dr. Beth L. Wellman and to Drs. Boyd R. McCandless and Charles C. Spiker for their valuable guidance in the conduct and analysis of this experiment. Appreciation is also expressed to the research assistants and Miss Eva Fillmore of the Research Station, and to the Iowa City Public Schools for their active cooperation.

This dissertation is filed as Publication No. 5476, University Microfilms, Ann Arbor, Michigan. Microfilm copy of complete manuscript of 126 pages, $1.58. Enlargements 6 by 8 in., 10 cents per page. Library of Congress card number MicA53-1609.

[2] The writer is now associated with the Department of Child Development and Family Relationships and the Agricultural Experiment Station at the University of Tennessee.

Hypothesis 1—The administration of the training program to kindergarten children will result in a statistically significant mean gain in the Total IQ's as obtained by the PMA.

Hypothesis 2—The experimental subjects (*Ss*) initially higher in mean PMA Total IQ will show significantly greater mean gains than those initially lower.

Hypothesis 3—The administration of the training program to kindergarten children will result in a statistically significant mean gain in the Full Scale IQ's as obtained by the WISC.

Hypothesis 4—The experimental *Ss* initially higher in mean WISC Full Scale IQ will show significantly greater mean gains than those initially lower.

Method

Training Materials

The training materials consisted of three main parts: (*a*) the teacher's manual, (*b*) group lesson charts, and (*c*) *The Red Book*. The latter was made up of 66 pictorial exercises each one of which centered around one or more of the various PMA factors. The group lesson charts were displayed to the group to demonstrate procedure in solving the various exercises. Lastly, the teacher's manual was principally concerned with exercise administration.

Subjects

The 107 *Ss* included in this study were members of four separate public school kindergarten groups in Iowa City, Iowa. The four groups as a unit consisted of 50 females and 57 males, the overall mean CA being five years, six months (5-6) with a range of 5-0 to 5-11. The CA's were computed from the date of the first test received by the child whether it was the PMA or WISC.

Only those *Ss* were used for whom both pre-training tests and post-training tests (hereafter, *pretest* and *posttest* respectively) afforded computation of the PMA Total and WISC Full Scale IQ's. Seventeen of an original 124 *Ss* were eliminated from final analyses because of failure to meet this criterion which resulted in a total sample of 107 *Ss*. Absences, nonrapport, and administrative errors were the principal reasons for incomplete records.

Experimental Design

A 2×2 factorial-type design was employed using intact classes as shown in Table 1. One of the two factors was classified as *treatment* which included experimental and control groups. The other factor, classified as *IQ level*, was composed of two parts, namely, an upper and a lower level

TABLE 1

FACTOR DESIGNATION AND NUMBER OF CASES FOR EACH OF FOUR GROUPS CONSTITUTING THE 2 x 2 FACTORIAL DESIGN TOGETHER WITH THE ASSIGNMENT OF THE TWO TEACHERS

Group	*IQ Level*	*Treatment*	*Teacher*	*N**
E_1	1	Experimental	1	28
C_1	1	Control	1	31
E_2	2	Experimental	2	25
C_2	2	Control	2	23

NOTE—The group designations in column 1, from the first to the last row, are to be read as follows: Experimental group, IQ Level 1; Control group, IQ Level 1; and so on, through IQ Level 2. The subscripts may also signify teachers. These symbols appear throughout the remainder of the report.

* Corrections for disproportionate cell numbers were used in the analysis.

(or Levels 1 and 2 respectively). The relative magnitudes of the PMA and WISC pretest mean IQ's constituted the criteria upon which the IQ levels distinction was made. These means are recorded in Table 2, Column 2, where it may be observed that in the case of the PMA, Groups E_1 and C_1 had means which were fairly high and relatively close, and that the means for Groups E_2 and C_2 were fairly close to one another but somewhat lower than those for E_1 and C_1. The same condition held true for the WISC pretest means, although the levels were separated by lesser IQ margins when compared to the PMA means. It may be added that Groups E_1 and C_1 arose from a generally *high* socio-economic population, and that Groups E_2 and C_2 arose from a generally *low* socio-economic population.

The use of experimental and control treatments was designed to provide a basis for testing Hypotheses 1 and 3. Similarly, the use of two IQ levels was designed to provide a basis for testing Hypotheses 2 and 4. In the case of testing the latter hypotheses, the *E* also planned to obtain additional information by computing Pearson *r*'s between pretest IQ's and IQ gains from pre- to posttests within the experimental groups.

The criterion scores used for the analyses of variance were the IQ *differences* (or gains) between the pre- and posttest results within each of the four groups for both the PMA and WISC separately. The analyses of both fourfold tables involved the same *S*s.

Experimental Procedure

The pre- and posttests were administered by a staff of research assistants from the Iowa Child Welfare Research Station. The staff underwent a one-month inservice training period concerned with test administration and

scoring prior to testing *S*s used in the present study. Two kindergarten groups, exclusive of the research groups, served as staff training *S*s.

All research *S*s were tested in their respective school buildings. None of the examiners was familiar with any *S* before testing. The pretests required a period of approximately six weeks to administer in the Fall of 1950. The posttests required about seven weeks to administer in the Spring of 1951. The WISC was administered individually to each subject,[3] whereas the PMA was given in two separate sessions to small groups containing approximately five *S*s each.

Although not possible in the *pre*testing program, examiner's personal schedules in the *post*testing program permitted: (*a*) control of test order such that approximately half the *S*s were assigned at random to one of the two combinations (PMA-WISC and WISC-PMA) within each group; (*b*) each PMA and/or WISC examiner to test an approximately equal number of *S*s in each group respectively; and (*c*) the random assignment of the given number of *S*s to each examiner. In general, *S*s were necessarily tested at the examiner's convenience during the *pre*test program. In certain instances, given pretest examiners did almost all the testing in a given school.

In conclusion, certain changes and additions were made with respect to both sets of test data in those instances where scoring and administrative criteria were indefinite. The changes and additions were accomplished through the generous assistance of the test authors and members of the project staff. Graphic extrapolation of some PMA scores was also made because of limited norms. A full description of the above operations may be obtained from the author or the original dissertation.

Training Program

The teachers participated in instructional periods at the Research Station concerning the use of the materials during the month prior to the onset of the training program. The program was administered to the two experimental groups during a 14-week period between the pre- and posttests in strict compliance with the suggested teaching procedures for the lessons in *The Red Book*. All experimental *S*s received the complete set of exercises, absences being accounted for by special makeup sessions.

Results

Table 2 includes the basic test results and components used to make the tests of significance. Two factorial-type analyses of variance were made of the pre- to posttest IQ *gains* (criterion scores) for both the PMA and WISC separately. The quadrantal table cell means and *SD*s of the gain distributions used in the PMA analysis of variance are given in the first four rows,

[3] Digit Span and Mazes were not administered as a part of the WISC.

TABLE 2

PMA AND WISC PRE- AND POSTTEST MEAN AND MEAN GAIN IQ'S WITH THEIR *SD*'S FOR EACH OF FOUR GROUPS

	Pretest		Posttest			
Group	*Mean*	*SD*	*Mean*	*SD*	*Mean Gain*	*SD of Gains*
			PMA			
E_1	109.57	16.51	121.89	13.85	12.32**	10.58
C_1	109.13	13.65	114.23	13.52	5.10*	9.80
E_2	87.88	13.38	101.20	15.28	13.32**	7.21
C_2	97.26	22.39	100.30	17.44	3.04	12.06
			WISC			
E_1	105.46	13.83	113.50	13.48	8.04**	8.52
C_1	101.48	9.05	108.94	10.44	7.45**	9.04
E_2	88.04	11.62	96.96	11.81	8.92**	7.67
C_2	94.57	16.19	104.13	16.98	9.57**	5.54

NOTE—See Table 5 for the corresponding significance ratios of the mean IQ gains.

* $p < .05$.

** $p < .01$.

Columns 6 and 7, of Table 2. Those used in the WISC analysis of variance are given in the lower four rows of the same columns.[4] Use of this design requires that groups be assigned at random to treatments within levels. Strictly speaking, this assumption was not met, since it was decided that a better estimate of training effects on children initially higher in PMA IQ versus those initially lower would obtain if the group having the higher pretest PMA mean in Level 1 (Group E_1) and the group having the lower PMA pretest mean in Level 2 (Group E_2) were used as experimental groups. This procedure was applied.

With respect to the other assumptions which underlie the use of this design, inspection of the PMA and WISC gain distributions suggested no reason to doubt the assumption of normality. Bartlett's method for testing the hypothesis of homogeneous cell population variance was computed for the PMA and WISC. Both chi-squares were nonsignificant, being between *p* values of .10 and .20.

[4] The statistical significance of each of these *individual group* mean IQ gains is reported at this point to avoid repetition later in text. Discussion of the rationale and procedures used in making these tests occurs subsequently in the results section. The analyses of variance now being described do not require information pertaining to the significance status of the individual group mean IQ gains even though the means are used in the 2×2 analyses of variance tables.

TABLE 3

SUMMARY TABLE OF THE 2×2 FACTORIAL ANALYSIS OF VARIANCE OF THE PRE- TO POSTTEST TOTAL IQ GAINS

Source of Variation	*df*	*ss*	*F*
Training (T)	1	1971.00	18.88**
IQ Levels (L)	1	42.86	< 1.00
Interaction ($T \times L$)	1	61.82	< 1.00
Within Cells (W)	103	10755.22	
Total	106	12788.64†	

† The *ss* of Total is not the sum of the *ss*'s in the same column because of the changes made in the *ss*'s of T, L, and $T \times L$ as a result of correcting for disproportionate cell numbers.

** $p < .01$.

A summary of the PMA analysis of variance is presented in Table 3. The within cells mean square (ms_W) was used as the error term for the T, L, and $T \times L$ effects. The only significant F ratio was that for the T category. The summary table for the WISC analysis (computed as in the case of the PMA) is not given, since all F ratios were less than unity, and therefore nonsignificant.

The tests of the main T effects for the PMA and WISC were relevant to Hypotheses 1 and 3 respectively. Similarly, the tests of the main L effects for the PMA and WISC pertained to Hypotheses 2 and 4 respectively. Since the tests of the main L effects were by no means unequivocal, Pearson r's were computed between pretest IQ's and IQ gains within the experimental groups for both the PMA and WISC. The r's are presented in Table 4. The only r among these which was significantly different from zero was the one obtained for PMA Group E_1.

Aside from testing the four major hypotheses, other tests of significance were performed which dealt with overall mean gains and individual group

TABLE 4

CORRELATIONS BETWEEN PRETEST IQ'S AND IQ GAINS WITHIN EACH EXPERIMENTAL GROUP FOR THE PMA AND WISC SEPARATELY

Group	*N*	*PMA*	*WISC*
E_1	28	—.55**	—.35
E_2	25	.01	—.30

NOTE—A test of the null hypothesis was made in each case where $df = N - 2$.

** $p < .01$.

mean gains. With respect to the former, a test of the null hypothesis was made in the case of the overall mean IQ gain of all *Ss* on the PMA and WISC separately. In each instance the overall mean gain was obtained from a pooled sample ($N = 107$) of the individual *S* gains. Then for a given distribution (PMA or WISC), a *t* test for related measures was computed. The variance estimate used in the given *t* test error term was the *ms* for "Total" which was obtained from the corresponding analysis of variance made earlier. The overall mean gains on the PMA and WISC were 9.47 and 8.40 IQ points respectively. Both means were significantly different from zero as recorded in the last row of Table 5.

TABLE 5

PMA AND WISC SIGNIFICANCE RATIOS OF THE OVERALL AND INDIVIDUAL GROUP MEAN IQ GAINS FROM THE PRE- TO POSTTEST PERIODS

Group	$df = N - 1$	*PMA*	*WISC*
E_1	27	6.25**	5.15**
C_1	30	2.73*	5.03**
E_2	24	6.37**	5.41**
C_2	22	1.39	5.56**
Overall	106	7.92**	10.77**

* $p < .05$.
** $p < .01$.

The significance of the difference between the mean gain for each group and zero was also tested by use of the *t* test for related measures. The group mean gains for both PMA and WISC were involved. The appropriate ms_W from the analyses of variance was used as the variance estimate in the various *t* tests. The selection of this variance estimate was based on the retention of the hypothesis for homogeneity of variance for the four group gain distributions on both intelligence tests. The mean gains for the individual groups for both PMA and WISC (numerators of the *t* tests) are given in Table 2, Column 6. Those values which were found to be statistically significant are designated there. Table 5 contains the significance *ratios* computed for both the overall and individual group mean gains.

DISCUSSION

Hypotheses 1 and 3 stated essentially that the experimental *Ss* would show significant mean gains on the PMA and WISC respectively as a function of training. The rejection of the null hypothesis with respect to the

criterion means of the main PMA *T* effects appeared to lend substantial support to the tenability of Hypothesis 1. On the other hand, the retention of the null hypothesis with respect to the criterion means of the main WISC *T* effects offered no noteworthy evidence as to the tenability of Hypothesis 3. These results might be accounted for by the fact that the PMA test items were strikingly similar to the items contained in the exercises of the training program. There appeared to be a marked dissimilarity between the training program items and those contained in the WISC which might have been related to the nonsignificant *T* effects obtained in the WISC analysis of variance.

It was interesting to note that in the instances of both the PMA and WISC, the overall mean gains were significantly different from zero, and furthermore, the same may be said about the mean gains within each school group with the exception of PMA Group C_2. The fact that some of the control groups showed significant gains[5] tended to uphold the general thesis that significant changes in scores on intelligence tests do occur. Under the present experimental conditions, however, one cannot definitely state the variables responsible for the changes in the control groups.

Hypotheses 2 and 4 stated that the experimental *Ss* initially higher in mean IQ would show significantly greater mean gains than those initially lower on the PMA and WISC respectively. The nonsignificance of both the PMA and WISC main *L* effects indicated that the obtained data were insufficient to justify the rejection or acceptance of either Hypothesis 2 or 4. The correlations between initial IQ and gains within each of the experimental groups for both tests afforded a better means for an attempted solution to the problem. Implicit in Hypotheses 2 and 4 was the prediction that such correlations would be positive and significantly different from zero. Inspection of the *r*'s obtained for the experimental groups presented in Table 4 definitely favored the rejection of Hypotheses 2 and 4, since all were either negative or near zero. Certainly there was no evidence to support the hypotheses. The correlational findings suggested an alternate hypothesis that perhaps the relationship between initial IQ standing and IQ gain is zero or negative.

The basic schema of the experimental design and its administration permitted the possibility that variation other than training might have in some systematic way affected the magnitudes of the criterion scores. With respect to the basic schema, inspection of Table 1 shows that groups were confounded with treatments and also with levels, since each cell contained a different group. The same table shows that teachers were confounded with levels but counterbalanced for treatments. The unfeasibility of completely controlling factors such as these constituted the major limitation of the research.

---

[5] The mean gain shown by Group C_2 in the WISC category was greater than for either of the experimental groups in the same category (see Table 2).

In regard to administration of the experimental design, it was indicated in the main text that circumstances precluded the possibility of achieving optimum conditions of, (*a*) test order and (*b*) examiner-to-subject and examiner-to-group testing assignment, in the *pre*testing program. However, this was not the case in the posttesting program.

Procedural steps were taken wherever practical to minimize the possible systematic effects of the factors mentioned above. The ways in which these factors were handled have been discussed in the experimental procedure section. The effects of these variables, if any, could not be obtained from the data. However, as a method of acquiring some evidence with respect to examiner differences, two analyses of variance were performed with the *post*test data—one for the PMA and the other for the WISC. The respective analyses were applied in each case to a two-way classification table, namely, *examiners x groups*. In both instances, the null hypothesis of no examiner differences was retained as were the two interaction ratios, where p was greater than .05 for all four F tests. These results provided valuable information in the sense that there was no explicit evidence to suggest that examiner differences had any systematic effect on the *post*test scores. One could not, however, justifiably make the same inference in the case of the *pre*test scores. Whether or not possible examiner differences exerted a systematic influence on the pretest scores is a matter of empirical fact which could not be determined from the test results.

Since it was possible that extraneous variables affected the criterion scores even though various ways of controlling most of them were applied, and since the assumption of the random assignment of groups to treatments within levels was not strictly met, an interpreter is bound by the principles of sound statistical inference and logic to be conservative in making generalizations from the obtained results. It is with these facts in mind that the following conclusions were formulated.

Summary and Conclusions

The purpose of this experiment was to test four hypotheses suggested by Thelma Gwinn Thurstone's kindergarten training program—*The Red Book*. Hypothesis 1 stated essentially that kindergarten *S*s who received the training program would show statistically significant mean Total IQ gains as obtained by the SRA Primary Mental Abilities-Primary (PMA). Hypothesis 2 stated that experimental *S*s initially higher in mean PMA Total IQ would show significantly greater mean gains than those initially lower. The same hypotheses (3 and 4 respectively) were made with regard to the Full Scale IQ as obtained by the WISC.

The 107 *S*s (57 females and 50 males) included in this study were enrolled in the kindergarten groups of each of four public schools in Iowa City, Iowa. The mean CA was five years, six months. A 2×2 factorial

experimental design was employed to test the hypotheses where the criterion scores were the pre- to posttraining IQ gains. Each cell of the fourfold table in question contained a different group. Analyses of variance were computed for the PMA and WISC gains separately. In each of these tables, the significance of the differences between the mean IQ gains for the control and experimental groups (treatment classification of the design) determined the final point of view adopted concerning the acceptability of either Hypothesis 1 or 3, as the case might be. *Low* and *high* IQ levels constituted the other classification of the factorial design, and they were differentiated and defined on the basis of the pretest IQ's. Then for a given table (PMA or WISC) the significance of the differences between the mean IQ gains for the low versus the high level was relevant, but not decisive, to the testing of Hypothesis 2 or 4, as the case might be. More germane to the testing of Hypotheses 2 and 4 were the correlations computed between initial IQ standing and IQ gains within each experimental group for both the PMA and WISC.

For each of the two fourfold tables of IQ gains, a statistical test was made to determine whether or not the overall (grand) mean differed significantly from zero. The same hypothesis was tested with respect to each of the cell (individual group) means in both analysis of variance tables.

The major limitation of the experiment was the unfeasibility of controlling completely for the possible effects of such factors as group differences associated with schools, examiner differences, teacher differences, and test order. However, steps were taken to minimize the possible effects of these factors. Whether or not such potential sources of variation affected the criterion scores systematically and significantly could not be determined from the data.

The following were the results obtained and conclusions made with respect to each of the four hypotheses and other relevant material. Where conclusions were founded upon statistical tests of significance, a coefficient of risk of $p = .05$, adopted prior to the experiment, was used.

1. From the PMA analysis of variance, it was found that the means of the IQ gains for the experimental treatment groups were significantly higher than the corresponding control groups who did not receive the training, thereby supporting Hypothesis 1.

2. The WISC analysis of variance yielded no statistically significant results, thereby precluding any judgment with respect to the confirmation or rejection of Hypothesis 3.

3. The *r*'s between pretest IQ standing and IQ gains within each experimental group were .01 and —.55 for the PMA, and —.30 and —.35 for the WISC. Thus, all were either negative or near zero, thereby casting considerable doubt on the tenability of Hypotheses 2 and 4. These findings suggested the alternate hypothesis that the relationship may be negative or zero where either intelligence test is used.

4. In the case of each intelligence test, the total sample ($N = 107$) gained significantly in mean IQ from the pre- to posttests. The PMA mean gain was 9.47 points, while the gains on the WISC produced a mean of 8.40.

5. The IQ gains from the pre- to posttests were statistically significant on both the PMA and WISC for each individual group, with the exception of one PMA control.

6. Results (4) and (5) appeared to uphold the general thesis that intelligence test scores do change significantly. However, it was not possible to specify definitely the factors determining the significant changes.

7. To the writer, the only significant meaning that may be associated with increases in the scores of the two tests, or any other such "intelligence" test, is the relationship between the gains and other operationally defined variables such as arithmetic or reading performance. The present study dealt only with questions concerning gains in IQ.

References

1. Thurstone, Thelma G. *Learning to think series. Teacher's manual for play and learn. The red book.* Chicago: Science Research Associates, 1948.

2. Thurstone, Thelma G. *Learning to think series. The red book.* Chicago: Science Research Associates, 1948.

3. Thurstone, Thelma G., and Thurstone, L. L. *SRA Primary Mental Abilities-Primary.* Chicago: Science Research Associates, 1946.

4. Thurstone, Thelma G., and Thurstone, L. L. *Examiner manual for the SRA Primary Mental Abilities-Primary.* Chicago: Science Research Associates, 1948.

5. Wechsler, David. *Wechsler Intelligence Scale for Children.* New York: Psychological Corp., 1949.

THE STATISTICS OF THE INDIVIDUAL CHILD: THE USE OF ANALYSIS OF VARIANCE WITH CHILD DEVELOPMENT DATA[1]

FRED N. KERLINGER
Wayne University

A number of methods have been devised to analyze the pattern of child development data, and these methods, to a certain extent, have lacked simplicity and precision. Without going into detail, it can be said that, as far as the data of the individual child are concerned, the methods are mainly graphical. Seriatim measurements of individuals' growth—in height, weight, intelligence, etc.—are expressed as the mean chronological age at which they appear in the standard population. These are called "growth ages" and are plotted against the individuals' own chronological ages (9). From an "age-unit" graph constructed in this manner the researcher is able to judge more or less subjectively whether a child's growth pattern is homogeneous (displaying a relative lack of dispersion among all the plotted growth ages) or heterogeneous (showing a relatively high degree of dispersion among the plotted growth ages). Certain hypotheses have been advanced as to the characteristics of homogeneous and heterogeneous growers and tentative predictions have been made as to the relative success in school of the two types of growers, as well as their relative degree of adjustment.[2] In short, much of the work has tried to get at the relationship, by one means or

[1] This study was made possible by the cooperation of Mr. Scott W. Street, Principal of the Best Elementary School, Ferndale, Michigan, and Mr. Roy E. Robinson, Superintendent of Schools, Ferndale, Michigan, who loaned the writer growth data to analyze, and who gave permission to publish the results of the analysis. Their kind help is gratefully acknowledged. Thanks are also due to Dean Willard C. Olson, School of Education, University of Michigan, under whom the writer studied child growth and development, and to Mr. Thomas Parsons, Dean Olson's Research Associate, who stimulated the writer to do the study and who gave him many ideas and suggestions.

[2] The work with which the writer is most familiar has been done at the Child Development Laboratories, University of Michigan, and this paper will be based on the approach used there.

another, between physical and mental growth, on the one hand, and school achievement and "adjustment," on the other hand.

As far as the writer knows, there have been few attempts made to use certain modern statistical ideas and techniques, particularly those of R. A. Fisher, to solve the problem of the analysis of individual development. That is, little effort seems to have been made to arrive at a "statistics of the individual child." Olson's organismic age (OA)—the average of various growth ages at a point in time—is one fruitful way of approaching the problem (7, 8). Other methods, too, have been devised, but the problem of homogeneity or heterogeneity of growth seems not to have been tackled in a clearcut and simple fashion.[3]

Purpose

The purpose of the present paper is to outline a method for getting at the "statistics of the individual child" in a manner quite different from previous methods. Specifically, growth data are analyzed to accomplish two main purposes: (1) to determine, in a precise and simple manner, the homogeneity or heterogeneity of a child's growth, and (2) to relate this method to the graphic age-unit method described above. Two methods are discussed: the usual parametric analysis of variance and a non-parametric method. Once this line of thinking and analysis is opened up it will be seen that other related techniques can be profitably used, depending on what kind of data the researcher is handling and what questions he wants to ask the data.

Method of Analysis

Child development researchers customarily measure children's height, weight, dentition, grip strength, wristbone development, reading ability, mental development, and so on (1, 4, 7). The Olsen-Hughes method of handling such measures is to convert them to comparable measures (age-units) in order to make comparisons and graphing possible and also to compute organismic ages (9). The present method of analysis uses Olson-Hughes comparable measures (age-units) and then, by applying simple analysis of variance to these measures, derives an index (the *F*-ratio) of homogeneity-heterogeneity. When the data warrant it, the significance of the difference between organismic ages at different points in time can

[3] Two of the methods pertinent to this paper are those of Hughes and Whaley (11). Hughes worked out a standard deviation of growth which is computed from the deviations of the measures of all growth ages at a single point in time from the mean of these measures (the organismic age). Whaley has criticized this measure for its inaccuracy due to the fact that the SD of fast-growing children tends to be greater than the SD of slow-growing children simply because of differences in slope (rates of growth). Accordingly, he has devised a trigonometric method for "correcting" standard deviations (11, pp. 32-38).

also be computed by extending the analysis in the manner usually used with correlated measures. In short, the significance of the differences between column (growth ages) and row (organismic ages) means is determined. Certain conclusions as to the child's growth can then be drawn, depending on the significance and relative size of the *F*-ratios.

The Ferndale, Michigan, public schools have been measuring children for about four years at approximately six-month intervals. The Olson-Hughes method of conversion and graphing has been used, and a large body of data is available. For the purposes of the present study, seven children's records were selected in one school. The principal was asked to pick children who were homogeneous and heterogeneous and fast and slow growers. The data of each child were then subjected to analysis of variance after conversion by a subtractive method. In addition, two graphs of the data of each child were drawn, one of the original data and one of the "subtractive" data. The graphs were used to validate the analysis of variance techniques, i.e., subjective impressions of homogeneity and heterogeneity were obtained from the graphs and the analysis of variance results checked against these impressions. Several non-parametric methods were also tried with varying degrees of success. It is believed that the method to be described constitutes an adequate statistics of the individual child, and that considerable clarification and understanding of the growth of the child may be derived from the use of the method. In addition, the mechanics of setting up the data and the computations themselves yield other useful information for the researcher. The rationale of the technique and its application are described below.

There are two important assumptions behind the present method of analysis. The first is that the child grows as a whole (7). It follows from this assumption that all the growth measures we get from the child—assuming the validity and reliablity of our measurements—are expressions of a "master plan" of growth. Statistically speaking, we can say that a number of measures of growth taken at a point in time are a sample from the same population, the overall growth of the child. We can further assume, for our analytical purpose, that if we take a number of measures at different points in time these measures are a random sample of the infinite number of measures that could be taken within a period of observation. With these two assumptions we have the basis for our technique. But there is a difficulty here. This is the fact that there will be correlation—and thus, seemingly, lack of randomness—between our various growth measures due to the simple fact of relatively uniform increases over time. However, if we establish a baseline, say the earliest age at which measurements are taken, and subtract from all subsequent measures taken at later points in time the number of months' difference between the later point and the baseline point, theoretically we would then have a more or less random sample of the growth ages of a child at different points in time. Actually, it makes

no difference where the baseline is taken; the ultimate statistical results will be the same (i.e., the *F*-ratios will be the same).

It must be clearly realized that the above discussion is theoretical. We know that the various growth ages differ, and we also know that the numerous measures we get are not, strictly speaking, a random sample. The point is that for purposes of analysis we make these assumptions and *then test them*. In other words, we assume that, without the differential effects of the growth ages—HA, WA, RA, etc.[4]—and by using the subtractive procedure, all the growth ages would be the same within the bounds of random error. Specifically, if we start at some baseline, say five years of chronological age (60 months), and reduce all the various growth ages to this baseline by subtracting the necessary number of months from each measured growth age at different points in time, say at 66, 72, 78, 84, etc. months, theoretically we would get the same figure throughout. (This also assumes, of course, that all ages have been converted to comparable ages (age-units) by the Olson-Hughes method.) That is, we should get, for chronological ages of 60, 66, 72, 78, etc., a growth age of 60 at all these points in time. And this is true for all the different growth ages. Of course we will never get the same converted "subtractive" growth age throughout. But we can assume regularity and can say that observed irregularities are random or chance errors. It is as though we were studying a parent population of "subtractive" height ages, say, and selecting a sample of these height ages. If height growth were quite regular, these ages would always be the same. But they are not, of course; a child will grow more in height during one six-month period than he will in another. Yet if we assume a pattern of growth—and this is an assumption that must be made—then the measured height ages should not differ from each other too much; they should form, if we were able to measure an unlimited number of these growth ages, a normal distribution clustered around a mean of, in this case, 60. Similar reasoning applies to all other growth ages.

Now, if we use the Olson-Hughes method of converting all growth ages to comparable age-units, and if we use the simple subtractive method described above, then we should have, theoretically, one large parent population of all growth ages from which our measured values of height, weight, mental ability, etc. are samples. Suppose we measured a child at six-month intervals for four years. We would have eight measures of height, eight measures of weight, eight measures of reading, and so forth. With the usual seven growth ages of the Olson method, we would then have a total of 56 measures. These 56 measures can be conceived of as a sample from one parent population of growth ages. In fact, this assumption is behind the Olson concept of organismic age and growth of the child as a whole (7). The fact that we know perfectly well that no child's growth ages—the

[4] HA = Height Age; WA = Weight Age; DA = Dental Age; GA = Grip Age; RA = Reading Age; MA = Mental Age.

various growths measured in different points in time—would be the same does not affect our reasoning seriously. We can say that observed discrepancies are the random or chance fluctuations that are to be expected from any measure.

In this paper we are mainly interested in the differences between the different growth ages—HA, WA, DA, GA, RA, and MA—of an individual child and the extent to which these are small enough to warrant our calling the child a "homogeneous grower" or large enough to justify calling him a "heterogeneous grower." More specifically, we ask: Is the child a "homogeneous grower" or is he a "heterogeneous grower?" How fast is his growth compared to some "normal" criterion? How slowly is he growing? What can be predicted about his school achievement, his ultimate height, weight, and so on? How about his adjustment? We are here mainly concerned only with the first of these questions. Put in another more precise way, we ask: Do the child's various growth ages differ significantly from each other? On the hypothesis of equality discussed above, they should not; they should be "like" a random sample. (They are not random because we measure the child at certain stated regular times and we have not selected these times, say, from a table of random numbers.) To make what is meant very clear, we should expect that, if the child was a perfect epitome of theoretical expectation, his height age at 60 chronological months would be 60 months; his height age at 96 chronological months would be 60 months (if we use the subtractive method); his dental age at 132 chronological months would be 60 months; his reading age at 84 chronological months would be 60 months. With the usual fluctuations of growth these "criterion" measures of 60 should be something like 58, 61, 62, 59, 64, and so on. And our assumption is that these fluctuations will be small within each growth age and between the various growth ages. Now, if *between* (among) the various growth ages the differences are small, as in the example given above, then we have a "homogeneous grower." But if the differences are large, for example at 72 chronological months we had: HA = 72; WA = 74; GA = 81; DA = 64; etc., and similar differences at other chronological ages, then we have a "heterogeneous grower," or, as Hughes calls it, a "split-grower."

Analysis of Growth Data

With the preceding theoretical discussion as a background, we can now examine an actual case. In Table 1 are the growth data of R.A., a child in the Ferndale school system who is a fast heterogeneous grower.[5] If the reader will plot these data, using five-month intervals, he will see this rather clearly. Two questions can be asked of the data: (1) Is the child a homo-

[5] Fictitious initials are used here and elsewhere to preserve the children's anonymity.

TABLE 1

ORIGINAL GROWTH AGES OF R.A.

Chronological Age	*HA*	*WA*	*DA*	*GA*	*RA*	*MA*
94	99	106	92	95	81	84
100	105	115	103	89	85	...
105	114	125	103	125	93	107
112	121	140	123	117	98	...
117	129	150	136	137	94	140
123	139	156	136	139	102	...
127	148	163	136	...	114	...

geneous or a heterogeneous grower? (2) If she is heterogeneous, how heterogeneous is she? The first step to be taken in answering these questions is to convert the data by the subtractive method described above. Table 2 gives the subtractive data, together with the computations necessary for the variance analysis. The baseline is 94 months. Table 3 shows the simple analysis of variance using the Within Groups (Columns) Variance as the error term. In most cases the analysis would go no further. Table 4 shows the analysis of variance in the case when the researcher

TABLE 2

SUBTRACTIVE GROWTH AGES OF R.A. WITH COMPUTATIONS NECESSARY FOR ANALYSIS OF VARIANCE

CA	*HA*	*WA*	*DA*	*GA*	*RA*	*MA*	Σ	*OA*
94	99	106	92	95	81	84	557	92.83
100	99	109	97	83	79	...	467	93.40
105	103	114	92	114	82	96	601	100.17
112	103	122	105	99	80	...	509	101.80
117	106	127	113	114	71	117	648	108.00
123	110	127	107	110	73	...	527	105.40
127	115	130	103	...	81	...	429	107.25
ΣX	735	835	709	615	547	207	$\Sigma X_t = 3{,}738$	
$(\Sigma X)^2$ ..	540,225	697,225	502,681	378,225	299,209	88,209	$(\Sigma X_t)^2 = 13{,}972{,}644$	
M	105.00	119.29	101.29	102.50	78.14	99.00	$M_t = 101.03$	
ΣX^2	77,381	100,155	72,189	63,807	42,857	29,961	$\Sigma X_t^2 = 386{,}350$	

TABLE 3

ANALYSIS OF VARIANCE OF GROWTH DATA OF R.A. (SIMPLE)

Source	*df*	*Mean Square*	*F*
Between Groups (Growth Ages	5	1227.16	14.77 (.001)
Within Groups	31	83.07	
Total	36		

wants also to test the significance of the difference between (among) the organismic ages. This is ordinarily not necessary since it seems that very seldom are such differences significant. With R.A., however, we do have a significant difference; the case was deliberately chosen to illustrate the method. Here, of course, we use the Residual Variance as the error term since we are also taking account of the variance due to rows.

R.A. is a heterogeneous grower. The *F*-ratio is 14.77 which is significant, for 5 and 31 degrees of freedom, at the .001 level. If an Olson-type graph is drawn the result is confirmed. It is even more apparent if a graph is drawn using the subtractive data. (In some cases the usual graph is clearer, but often graphs drawn from the subtractive data are more satisfactory.) R.A.'s *F*-ratio ranks third among the seven children. That is, if we have approximately the same degrees of freedom for entering the *F*-table—and the degrees of freedom for these seven cases are more or less alike since with so many degrees of freedom the changes in the *F* entries are relatively small—we can legitimately compare the ratios of different children. Table 5 lists the *F*-ratios of the seven children. Graphs of both kinds were made for all the data and subjective evaluations from the graphs of homogeneity-heterogeneity confirm the *F*-ratios. It can be seen that the first five children are heterogeneous growers—all *F*-ratios are significant at the .001 level. The last two children are homogeneous growers. The .01 level was set for

TABLE 4

ANALYSIS OF VARIANCE OF GROWTH DATA OF R.A.
(FOR CORRELATED MEASURES)

Source	*df*	*Mean Square*	*F*
Between Columns (Growth Ages) ..	5	1227.16	23.04 (.001)
Between Rows (Organismic Ages) ..	6	207.23	3.89 (.01)
Residual: C x R	25	53.27	
Total	36		

TABLE 5

F-RATIOS AND *H* VALUES OF CHILD DEVELOPMENT DATA FOR SEVEN FERNDALE SCHOOL CHILDREN

Individual	*F*	*H*
S.T.	18.73 (.001)	23.37 (.001)
D.A.	16.08 (.001)	23.78 (.001)
R.A.	14.77 (.001)	23.22 (.001)
V.B.	10.06 (.001)	24.34 (.001)
S.K.	8.27 (.001)	19.80 (.01)
L.D.	2.16 (n.s.)	15.74 (.01)
B.E.	1.77 (n.s.)	11.23 (.05 — n.s.)

significance. (This is an arbitrary choice, but one which seems to agree well with the data and the graphs.)

The difference among R.A.'s organismic ages is also significant, if not at as high a level—$F = 3.65$ (.01). Note, too, that the ratio itself is not nearly as large as the *F* for Between Columns (Growth Ages). This means, if we inspect the means of the rows (the organismic ages), that R.A. has a definite tendency to be an accelerated grower, i.e., the slope of her growth is steeper than that of other children. Notice how the OA's increase steadily from 92.83 to 107.25 with one mean, 108, not in rank order. Theoretically, they should all cluster around 94.

Table 5 also includes the *H* statistics computed from the data of the seven children. It will be noted that the *H* statistic agrees fairly well with the *F* statistic, except in the last two cases. These results seem to be typical. The writer worked a good deal with other child growth data while trying to perfect the method, and the only statistic which seemed to be consistent with the *F* statistic was *H*. But there are times, of course, when any non-parametric method will not agree with its parametric companion. Generally speaking, *H* seems to be well suited to these data, especially when they are incomplete, i.e., when some one or the other of the growth measures was not taken at one of the points in time. This could be interpreted as a weakness of the method since researchers often do not get all their growth measures of one child at the same time. However, when this is the case, interpolation can be resorted to. The author has tried it with other data and it seems to work satisfactorily. A description of the *H* statistic can be found in Walker and Lev (10). A more complete discussion by its originators, Kruskal and Wallis, can be found in (6). Suffice it to say here that it is a rank-order statistic which yields a measure approximately equal to chi-square. The writer also tried Friedman's χ^2_r (3) and Kendall's *W* (5), but these were unsatisfactory because of missing data in some of the growth

ages. It should be noted that, with this type of data, the non-parametric statistics seem not to be "sensitive" to homogeneous growers. Or it might be better to say that they are usually "oversensitive," they usually show significant differences when there are none.

A further refinement of the method should be mentioned. If a set of data yields a significant *F*-ratio, it may in some cases be profitable to inspect the means of the growth ages to discover if any one particular growth is causing the high *F*. Should this be the case, a new analysis of variance can be computed omitting this very discrepant growth. (This can easily and quickly be done by subtraction in the original table, and then by re-computation of the sums of squares, the variances, and the *F*-ratio.) If the new *F* value is not significant, or is greatly reduced, then we may know that the child is a homogeneous grower except for the one discrepant growth. (This seems to happen quite often with Reading Age.) On the other hand, if the new *F* value is still very significant, or does not reduce too much, we can conclude that the child is really a heterogeneous grower. R.A.'s case is also a good illustration of this. She has a much lower reading age than her other ages. Is this low RA causing the significant *F*? The analysis was done without RA, and *F* is 4.10, which is not significant. (It is significant at the .05 level, but the criterion set up was the .01 level.) Thus, if we discount RA, R.A. is a homogeneous child. Actually, because there is correlation between the rows, we could get a more refined error term, and the *F*-ratio would probably be significant. But this is sufficient to illustrate the point. Besides, with these data, it is probably best to err on the conservative side. In other words, this addition to the technique enables the researcher to be quite precise about his estimate of homogeneity and heterogeneity. Similar reasoning would of course apply to the analysis of organismic ages (rows).

Conclusion

In general, then, analysis of variance can be used with child growth data to determine the significance of the difference between growth ages and between organismic ages. In most cases, apparently, it is unnecessary to test organismic ages since their differences at different points in time are not often significant, certainly not nearly as significant as the differences between the various growth ages. This finding, if confirmed by further analysis of this kind, should be of value in the field of child development, especially in view of Olson's thinking on the use of organismic age as a predictive measure because of, among other things, its stability (7, pp. 180-181). Significant differences in organismic ages would indicate one or two of three things: (1) a rapidly accelerating grower; (2) a rapidly decelerating grower; (3) an erratic grower. Among the seven Ferndale children we have one or two rapid accelerators.

It is not possible to say now how useful this method of analysis can be. For example, it takes about forty minutes (or less if a number of children's data are analyzed and prepared forms are used) to record and compute the data for one child. Is this too much time when one has many children's data to analyze? The method would not seem to be practicable for schools, especially if a modern computing machine is not available, but for the research worker it would seem to be a valuable tool not only for helping to categorize individual children but also for making comparisons, seeking relationships, and testing hypotheses. Moreover, it gets away from the older almost sole dependence upon individual analysts' unstandardized interpretations. Within certain limitations the method not only clearly distinguishes homogeneity from heterogeneity but also gives the degree of homogeneity and heterogeneity. It is easy and relatively quick to use and supplies a precise, usually unambiguous statistic for interpreting the growth of an individual child. And as an artifact of the method the researcher has a data table which makes the study of a child's total growth as well as his individual growths feasible. Certain measures of relationship, e.g., Kendall's W, can also be used with the subtractive data. Naturally, the method is not perfect. At this writing, for instance, the ages of about six to ten would seem best for this type of analysis. It may be that the method would break down, or have to be considerably altered (although the basic idea would remain the same, of course), perhaps, if data from children who had attained puberty were analyzed. Then, too, there must be homogeneity of variance if the method is to be valid. And, unfortunately, one very discrepant measure, say a very low grip strength measure at age seven, can cause heterogeneity of variance. Of course this can be easily detected and corrected by the researcher. (S.T. was such a case. The removal of the one very discrepant measure, however, rectified the situation. It is also important to note that this measure may have been a mistake of some kind.) Bartlett's test for homogeneity of variance was applied to the seven sets of data and in all cases except the one noted above the variances were homogeneous. The researcher should be wary, however; this may not always be the case. (If it is found that the variances are heterogeneous, it may be possible to use the $\log X$ transformation (2). It should also be noted that finding significant differences in variances might in itself be very useful information for the child researcher, e.g., it might indicate unusual differences in rates, inadequate measurement, or some sort of aberration.

Once the basic idea of the method is grasped there are other applications and methods that might be worked out. For example, analysis of covariance could easily be applied to twin measurements and would yield not only F-ratios of value but also different types of coefficients of correlation (r for Between Groups and r for Within Groups, the latter being a more accurate estimate of the population r). The dualistic-organismic controversy might also be approached with the method (11). Differences in rates of growth

can also be tested perhaps by subtracting growth ages from each other at different points in time and then by using the remainders as criterion measures. But before the true usefulness and validity of the method can be adequately determined it should be used with a fairly large representative sample of children, preferably to test specific hypotheses in which homogeneity and heterogeneity and perhaps organismic ages are variables. At any rate, the least that can be said is that it *is* an approach to a statistics of the individual child and it does seem to supply a rather precise index for judging the homogeneity and heterogeneity of child growth and for judging the stability of a child's growth as measured by organismic age.

References

1. Bayley, Nancy, and Carter, H. Child development—V. Physical growth. In W. S. Monroe (Ed.), *Encyclopedia of educational research.* (Rev. Ed.) New York: Macmillan, 1950.
2. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
3. Friedman, M. The use of ranks to avoid the assumption or normality implicit in the analysis of variance. *J. Amer. statist. Ass.,* 1937, 32, 675-701.
4. Jensen, K. Physical growth. *Rev. educ. Res.,* 1952, 22, 391-420.
5. Kendall, G. *Rank correlation methods.* London: Charles Griffin, 1948.
6. Kruskal, W. H., and Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. statist. Ass.,* 1952, 47, 583-621.
7. Olson, W. C. *Child development.* Boston: Heath, 1949.
8. Olson, W. C., and Hughes, B. O. The concept of organismic age. *J. educ. Res.,* 1942, 36, 525-527.
9. Olson, W. C., and Hughes, B. O. *Manual for the description of growth in age units.* Ann Arbor: Child Development Laboratories, Univer. Elementary School, Univer. of Michigan, 1950.
10. Walker, Helen M., and Lev, J. *Statistical inference.* New York: Holt, 1953.
11. Whaley, F. L. A study of the relationship of the development of the child as a whole at the elementary level to high school achievement and activity participation. Unpublished doctor's dissertation, Univer. of Michigan, 1953.

A DEVELOPMENTAL STUDY OF SYMBOLIC BEHAVIOR[1]

ALLEN HODGES
University of Tennessee[2]

The double alternation problem has long been recognized as a measure of symbolic behavior in the animal laboratory and has been used in tracing phylogenetic development (1, 4, 5, 7, 8).

The value of the double alternation technique with human subjects has not been overlooked (2, 3, 6, 9); however, a search of the literature reveals no attempt to include a representative number of normal subjects from which an estimate of maturation of symbolic behavior in normal children might be obtained.

The purpose of this study, therefore, is to trace the ontogenetic development of symbolic behavior in normal children, exerting necessary controls in order to obtain as pure a measure as possible of the effect of chronological age on double alternation behavior.

METHOD AND PROCEDURE

Selection of Subjects

All subjects were selected from a normal public school population.[3] The procedure followed in selecting subjects was to select randomly 40 male pupils from each of the six age groups included, 10 subjects having an IQ between 80 and 90; 10 having IQ's from 90 to 100; 10 having IQ's from 100 to 110; and 10 subjects with IQ's between 110 and 120. All IQ's were obtained by means of the Primary Mental Abilities Test (10).

Table 1 contains the chronological and composite mental age characteristics of the six age groups.

---

[1] The author is indebted to Drs. G. R. Pascal and E. O. Milton of the University of Tennessee for the advice and guidance given in completing the dissertation upon which this article is based.

[2] Now at Southern Minnesota Mental Health Center, Albert Lea, Minnesota.

[3] Gratitude is expressed to Mr. Hilary Parker, Superintendent of the Oak Ridge, Tennessee, Public Schools, for his cooperation in obtaining subjects.

TABLE 1

CHRONOLOGICAL AND MENTAL AGES OF THE SUBJECTS

Age Group	N	Chronological Age		Mental Age	
		Mean	*SD*	*Mean*	*SD*
6 to 7	40	6.54	.31	6.21	.84
7 to 8	40	7.50	.28	7.59	.93
8 to 9	40	8.41	.34	8.15	.77
9 to 10	40	9.50	.39	9.25	.81
10 to 11	40	10.39	.36	10.41	.63
11 to 12	40	11.41	.24	11.42	.69

In addition to the criteria already mentioned, an attempt was made to limit the possible effect of socio-economic status by selecting subjects from three school districts in which family earning power, educational level of parents, and type of housing were very similar.

PROCEDURE

After selecting the 240 subjects in the manner described, the double alternation problem was presented to each subject individually. This particular version of the double alternation problem was first used by J. McV. Hunt in studying the aftermath of a partial bilateral frontal lobectomy (3).

The Card Test is simply administered, requiring five ordinary playing cards. Four of the cards are black (spades or clubs) while one is red (diamonds or hearts). The red card is designated as the goal card. The cards are shuffled by the examiner and presented face down before the subject. The goal card is placed twice on the left-most end and twice on the right-most end according to the double alternation sequence. The subject is simply instructed to find the red card. A complete trial consists of the LLRR sequence and the test is continued for thirty trials or until the subject is able to obtain two successive correct trials (LLRR, LLRR).[4]

RESULTS

A biserial coefficient of correlation for success vs. failure and mental age yields a coefficient of $.59 \pm .05$. When success vs. failure and chronological age are correlated, a biserial r of $.77 \pm .03$ is obtained. These results are in keeping with the findings previously reported by Gellermann (2) and Hunter and Bartlett (6) in which both CA and MA appear related to double alternation behavior.

[4] In a forthcoming article, Dr. J. McV. Hunt will present a more detailed description of the Card Test. Appreciation is expressed to him for allowing his unpublished manuscript to be read and referred to in this paper.

In order to obtain an estimate of the growth pattern for symbolic behavior as measured by this technique, the average percentages of solution as a function of chronological age were plotted for the 240 subjects. The results are contained in Figure 1. The results show an almost rectilinear progression of success with increased chronological age.

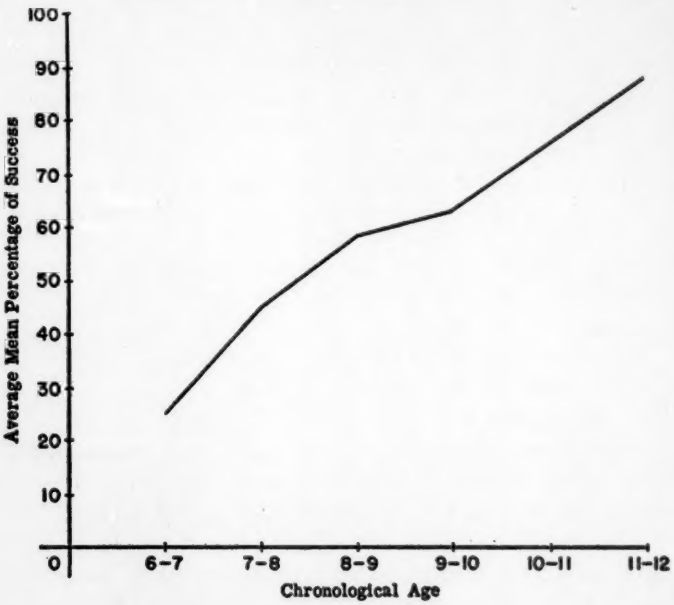

FIGURE 1—Average Mean Percentages of Success on the Double Alternation Problem as a Function of Chronological Age.

This estimated growth curve of the development of symbolic behavior in normal children between the chronological ages of 6 and 12 years has theoretical importance for child development as well as for the psychology of thinking. These results suggest that symbolic behavior follows a growth pattern similar to other patterns of biological growth.

SUMMARY

Two hundred and forty male children between the ages of 6 and 12 years were randomly selected from a normal public school population. A variation of the double alternation problem was administered to each

subject individually. Biserial coefficients for success vs. failure and mental age and chronological age were found to be .59±.05 and .77±.03 respectively.

An approximation of a growth curve of symbolic behavior was derived which appears to follow a growth process similar to other biological phenomena.

Bibliography

1. Gellermann, L. W. The double alternation problem: I. The behavior of monkeys in a double alternation temporal maze. *J. genet. Psychol.*, 1931, 39, 50-72.
2. Gellermann, L. W. The double alternation problem: II. The behavior of children and human adults in a double alternation temporal maze. *J. genet. Psychol.*, 1931, 39, 197-226.
3. Hunt, J. McV. A case of partial bilateral frontal lobectomy. *Amer. J. Psychiat.*, 1940, 96, 1063-1083.
4. Hunter, W. S. The temporal maze and kinesthetic sensory processes in the white rat. *Psychobiology*, 1920, 2, 1-18.
5. Hunter, W. S. The behavior of raccoons in a double alternation temporal maze. *J. genet. Psychol.*, 1928, 35, 374-388.
6. Hunter, W. S., and Bartlett, S. C. Double alternation behavior in young children. *J. exp. Psychol.*, 1948, 38, 558-567.
7. Hunter, W. S., and Hall, B. E. Double alternation behavior of the white rat in a spatial maze. *J. comp. Psychol.*, 1941, 32, 253-266.
8. Hunter, W. S., and Nagge, J. W. The white rat and the double alternation temporal maze. *J. genet. Psychol.*, 1931, 39, 303-319.
9. Stolurow, L. M., and Pascal, G. R. Double alternation behavior in mental defectives. Paper presented to Division of Experimental Psychology, American Psychological Association, Pennsylvania State College, 1950.
10. Thurstone, L. L. *The Primary Mental Abilities Test.* New York: Science Research Associates, 1948.

AN EXPERIMENTAL TEST OF THE ACCELERATED LONGITUDINAL APPROACH

RICHARD Q. BELL
National Institute of Mental Health

Data from the third Harvard Growth Study which show a prepubertal growth spurt have been offered in the past as support for the necessity of longitudinal studies, since it has been maintained that the averaging of individual curves in a cross-sectional study would have concealed the phenomenon (1, pp. 269-270). The data in question have been analyzed by another method to test the hypothesis that a customary longitudinal study was not necessary to establish the existence of the growth spurt.

ACCELERATED LONGITUDINAL METHODS

The method termed "convergence" has been advocated as a means of meeting some research needs not satisfied by either a cross-sectional or longitudinal approach (2). This method consists of making limited remeasurements of cross-sectional groups so that temporally overlapping measurements of older and younger subjects are provided. The remeasurements may be used merely as a way of determining whether trends which would otherwise be seen only between different age groups are corroborated within short time periods for each age group (5). The method may also provide a means of actually linking up individuals or sub-groups between adjacent segments of a developmental curve, each segment consisting of a limited longitudinal study on a different age group.

TESTING THE ACCELERATED APPROACH

To gain some notion of the problems encountered in the latter type of application, physical growth data obtained from the third Harvard Growth Study (4) were treated as though they had been obtained in two temporal segments on different age groups. Growth curves constructed from different segments could then be compared with curves constructed from the measurements of single individuals over the entire time span involved.

From a group of 125 female subjects on whom physical measurements were obtained at times near half-yearly anniversaries, 82 cases were selected who were measured in the main within a month and a half of half-yearly anniversaries. This was necessary to eliminate the possibility of attempting to link apparently overlapping measurements on individuals actually measured at widely different times. The group of 82 was stratified on the basis of types of standing height growth curves, and individuals from each

stratum assigned at random to one of two groups. The result was two roughly comparable groups for whom equivalent measurements were available, each group consisting of 41 cases. Measurements of standing height at ages 10.5, 11.5, 12.5 and 13.5 were available for each group of girls. This was the critical portion of the total growth curve and was the only portion selected for the present study. Other portions would also be studied in an actual field test.

To test the extent of error introduced by the accelerated approach one group of 41 cases, termed the first segment group, was treated as though it had only been measured at ages 10.5, 11.5, and 12.5. The second segment group was treated as though it had only been measured at 11.5, 12.5, and 13.5. The problem was to match individuals in the first segment group with those in the second group on the basis of standing height measurements at ages 11.5 and 12.5. A composite curve for each "individual" could then be constructed from the 10.5, and 11.5 measurements of one individual and the 12.5, and 13.5 measurements of another. The composite curve could be compared with the actual curve for one of the individuals used in the composite, since in actuality measurements are available from 10.5 through 13.5 for all individuals.

Matching Procedure

Each individual in the first segment group was matched with an individual in the second segment group for whom the smallest "D" measure could be obtained by getting the square root of the sum of the squared differences between measurements on standing height at 11.5 and 12.5. This is an application of D technique to correlated measurements (3). Measurements were expressed in centimeters and were not converted into standard scores. The matching procedure can be expected to capitalize on chance similarities between individuals who are in reality different. There may also be some question concerning the method of applying D. In this instance however a check is available on whether the matching is effective and the methods may be judged at least in part by the results.

Results

Figure 1 presents critical portions of curves taken from the original statistical treatment by Shuttleworth (6). Figure 2 is a presentation of the curves obtained by the accelerated method. Both figures are shown in terms of increments in standing height obtained by getting differences between measurements on each half-anniversary of measurement.

Curves plotted with open circle points in Figure 2 represent for three menarcheal groups composite curves constructed from the linked measurements of matched individuals. Menarcheal group is determined on the basis of information which would have been available at age 13-7 on the

individual from the second segment group with whom the individual from the first segment group was matched. Curves plotted with solid circles represent increments for actual measurements of single individuals, averaged by menarcheal groups. In the latter case the menarcheal group assignment is again in terms of the menarcheal age of the individual in the second segment group with whom the individual from the first segment group was matched. Curves plotted with solid circles are in effect average increment curves for the second segment group from age 10.5 through 13.5. Thus Figure 2 affords a comparison of actual and reconstructed growth curves.

The pre-pubertal growth spurt is still apparent in the data yielded by the accelerated approach. A statistical test is not provided since the phenomenon of a growth spurt was apparently noted originally on the basis of inspection of the growth curves, and seems available on the same basis in the present data.

Discussion

The results of the present limited test of an accelerated longitudinal approach suggest at least in this instance that one basic phenomenon derived from a full term longitudinal study could have been noted in a more limited study. The present test was limited to reconstructing only one portion of a developmental curve. Only one year's time and one measurement on the experimental population would have been saved by this procedure. The present test is assumed to be primarily informative on the feasibility of linking curves from different individuals and only secondarily representative of what an actual field test of such a procedure would involve. It is apparent that one group might have been selected and measured at age 10.5 and measured again at 12.5 and 13.5. The other group might have been selected and measured at age 12.5 and measured again at 13.5 and 15.5. At the same time that this portion of the curve was being explored two older age groups could be selected so as to reconstruct a segment of the developmental curve at a later point in time. A ten year span could be approximated in three years by this system.

The accelerated approach tested above is suggested as an alternative developmental method. It is obvious that it will introduce considerable error into data collected, and that in some instances it might produce quite misleading results. For example, if in the basic phenomenon under investigation similarities between individuals at one point in time are followed by unpredictable differentiation at a later point, the matching of cases at the earlier point in time would produce quite confusing results. On the other hand, it seems quite difficult to conceive of a developmental process which would not provide at some points the means for prediction of later changes. It would seem in most instances that some difference in level, shape, or dispersion of measurements in a curve would be predictive of later differentiation.

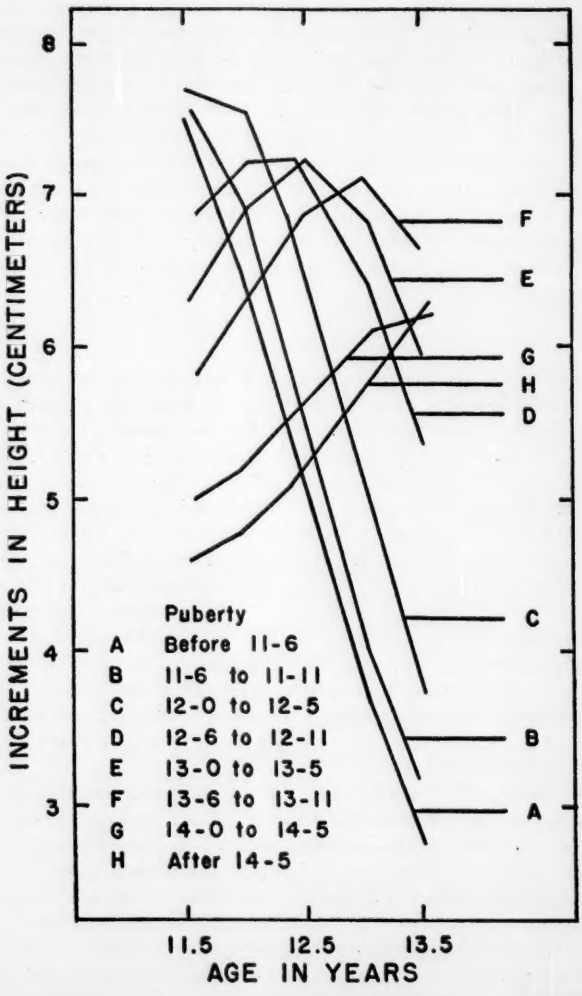

FIGURE 1—Average Annual Increments in Standing Height of Eight Groups of Girls Reaching Puberty at Different Ages. (From Shuttleworth, 6, p. 32)

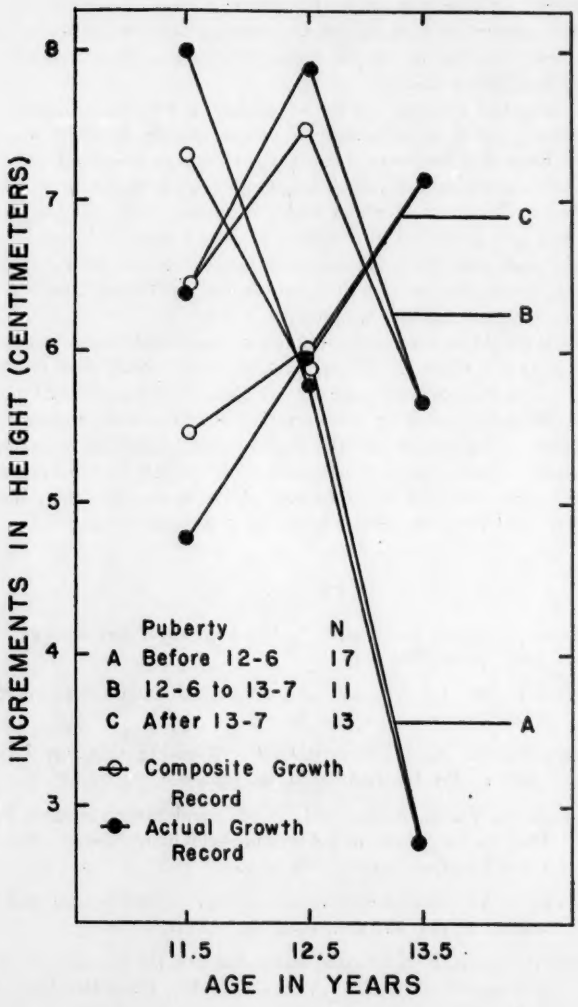

FIGURE 2—Average Annual Increments in Standing Height for Actual and Composite Growth Records.

Another problem is that relatively large samples are required for the purpose of matching cases. A second segment sample should be large enough so that close matches for first segment cases are available. In general measurement error introduced by matching cases or linking sub-populations must be offset by samples typically larger than those usable in customary longitudinal studies.

It is suggested that one use for an accelerated longitudinal approach is in obtaining initial approximations of results in order to make possible a sharper focus in a longitudinal study. Leads may be developed which the more time consuming and difficult longitudinal approach may be specifically designed to illuminate. Without such preliminary leads the longitudinal study may not be efficient in providing data for critical problems. A cross-sectional study may also be used for such purposes but, as has been pointed out (2), it provides no way of approximating individual curves nor of checking sampling between age groups.

Since it should be much more difficult to match individuals successfully on the basis of a single variable, as was done in this study, than in a situation where several measurements are available on each individual, tests of the accelerated approach in a multivariate situation seem warranted. In particular, applications to more global personality data should be attempted. More rapid approximations of longitudinal results may be most needed in this area, since social changes over a long period of time alter the value and implications of studies of culturally anchored personality variables.

References

1. Anastasi, Anne, and Foley, J. P. *Differential psychology.* New York: Macmillan, 1949.

2. Bell, R. Q. Convergence: An accelerated longitudinal approach. *Child Develpm.*, 1953, 24, 145-152.

3. Cronbach, L. J., and Gleser, G. C. Assessing similarity between profiles. *Psychol. Bull.*, 1953, 50, 456-476.

4. Dearborn, W. E., Rothney, J. W. M., and Shuttleworth, F. K. Data on the growth of public school children. *Monogr. Soc. Res. Child Develpm.*, 1938, 3, No. 1.

5. Miles, C. C. Influence of speed and age on intelligence scores of adults. *J. gen. Psychol.*, 1934, 10, 208-210.

6. Shuttleworth, F. K. Sexual maturation and the physical growth of girls age six to nineteen. *Monogr. Soc. Res. Child Develpm.*, 1937, 2, No. 5.

THE EFFECTS OF A MOTOR HANDICAP ON PERSONALITY: II. THE EFFECTS ON INTEGRATIVE ABILITY[1]

CHARLES WENAR
University of Illinois, College of Medicine

The term "integrative ability" will be used in the present study to include the following components: (*a*) Complexity—i.e., the number and variety of ideas which a child can bring to an unstructured task; and (*b*) Organization—i.e., the ordering of separate elements or ideas into coherent units. Experimentally these components will be operationally defined in terms of the task used in the present study. Clinically they can be regarded as an index to the resourcefulness the child brings to the everyday problems of living. The emphasis here is on conscious choice, planning, reasoning and conceptualization, and the central interest is in how well the child can maintain that combination of intellectual freedom and control which is regarded as a sign of mental well being.

There are two theoretical reasons for predicting that a motor handicap would have an adverse effect on integrative ability. A child with a motor handicap may well be more vulnerable to anxiety, and one of the most pronounced effects of anxiety is a decrease in the ability to deal with new and complex situations in a constructive way. This, in turn, should increase the possibility of unhealthy or deviate solutions to problem situations such as constriction, impulsivity, or fragmentation. Second, a motor handicap limits the range of experience of the child and one can infer that this, in turn, can limit the richness of the ideation which he brings to a task.

The following experiment was designed to test the hypothesis that a motor handicap produces a decrease in integrative ability and corresponding increase in deviate solutions to problems.

METHOD

Subjects

The same subjects were used in this study as in a previous one (2): 12 children with no motor handicap, 12 with a mild motor handicap, and 12 with a severe motor handicap. All were pupils in the Chicago public school

[1] These studies were sponsored by The Coordinated Program for Handicapped Children and were done at The Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago, Illinois.

system, between the ages of 8 and 10, of average or above intelligence, with no other severe physical handicap.

Experimental Task

For both theoretical and practical reasons the Bolgar modification of the World Test (1) was selected as being the most suitable instrument for measuring integrative ability. This test consists of 200 small, attractive toys representing houses, trees, animals, people, etc. Because of this wealth of objects the test can be used to tap the ingenuity and flexibility of the child's thinking. In addition, the realistic nature of the material places demands on the child which highlight the kind of logic and the maturity of conceptualization which he can bring to a problem. Thus it seemed admirably suited to reveal the factors of complexity and organization which constitute the definition of integrative ability.

In regard to practical considerations, the test requires the minimum amount of verbal communication and the material, with only slight modifications, can be handled by all but the most severely handicapped. These modifications consisted in mounting the animal figures and fences on cardboard and the human figures on round lead weights to give them stability; presenting the pieces in an upright position in random order instead of in boxes; and the experimenter's offering to place pieces at the child's direction during construction in the case of the severely handicapped individual. In general, it was felt that if the child had an idea about the material he managed to get it expressed the way he wanted it expressed.

The instructions were to "make something" with the pieces, and the emphasis was on doing anything the child wished and taking as long as he wanted. At the end of the task a general question was asked, "What have you made?" In addition the question "What is this?" was asked about specific units of the child's World which he did not identify.

Scoring

For the present study the usual scoring method of counting the number of objects or the number of categories of objects was considered as inadequate for two reasons; first, the emphasis on sheer quantity put the severely handicapped children at a disadvantage because of their difficulty in manipulating large numbers of objects; and, second, such a scoring system seemed to have little direct relation to integrative ability as defined above. Therefore a new scoring system was devised.

As a first step all the Worlds were examined and the distinct kinds of component parts were recorded. Fifteen classes or units of activity emerged: human activities, animal activities, living units, water scenes, fire scenes, train scenes, public buildings, blocks, traffic scenes, farm scenes, zoo scenes, airports, animal pens, parades, and global identifications. Then within each unit of activity, the constructions were rated along the dimension of sim-

plicity to complexity. There were two criteria used in making such ratings, the structural aspect of the construction or what the child actually built, and the conceptual aspect or what the child verbalized about the construction during the enquiry.

The actual assigning of scores within each class followed the theoretical notion that the simplest unit should have the lowest score and that the numerical value should increase as the unit became more complex. Thus the mere placing of an object on the table without elaborating it or differentiating it in any way was given no score; likewise the mere identification of an object ("This is a church") was not scored since there seemed to be no integrative activity involved. The general rule was that the lowest score was given to the unit in which two classes of objects were related (a house with a tree), or the simplest of differentiations was made ("This is the Chief of Police's car"). The score increased as a greater variety of objects were brought into meaningful relation with one another. In all cases higher scores were given only when new categories of objects were related, not when more of the same class were added; e.g., a house with three trees was scored no higher than a house with one tree, but a house, a tree and a fence was scored higher. Thus quantity of objects was relatively unimportant. The highest score always represented the most complex and unified construction in a particular class. For example, in the human activities class a score of 1 was given to such simple descriptions as "A lady selling food," whereas a score of 5 was given to a complex unit made up of at least four classes of related objects, such as "A lady coming to shut the gate after the dogs got the animals back in the pen."

In addition to this detailed analytic scoring system, a global rating of the Worlds was made which could serve as another kind of measure of integrative ability. Such a rating served two purposes. The first was to discover whether there was an overall pattern characteristic of the handicapped child which might be obscured when only the elements were scored. The second was to provide a more sensitive measure of deviate thinking since the analytic scoring proved to be inadequate in this respect.

An examination of the Worlds revealed that they fell into three groups, one of which could be considered as a healthy mode of dealing with the task, and the other two as representing deviations from health. First there were the well integrated Worlds (I), whose total pattern was characterized by a richness of ideation and a synthesis of elements. Then, on the one hand, there were Worlds which became progressively constricted (Cs) with a decreasing number of activities and increasing impoverishment of structure. On the other hand, there were Worlds characterized by an overinclusiveness (O) in which the child could not limit his construction but included extraneous and unrelated objects in a perseverative or jumbled manner. Using such a rating scheme, 9 Worlds were judged Integrated, 16 as Constricted, and 11 as Overinclusive. Using such a scheme, one could

not only test the hypothesis that a motor handicap produced a significantly greater number of deviate solutions to the task, but also whether severity of handicap was associated with a particular kind of deviate solution.

However, it might also be argued that the difference between handicapped and non-handicapped children may not be a clear cut one of healthy and deviate solutions, but merely an exaggeration in the handicapped groups of the kinds of deviate solutions found in the non-handicapped children —i.e., the differences in their Worlds may be one of degree instead of one of kind. In order to test this hypothesis also, a second method was devised. The extremes of the Cs and O Worlds, representing strikingly impoverished or chaotic Worlds, were placed in a Non-Integrated group. The remaining constructions, representing a transition from extreme deviation to good integration were placed in a Moderately Integrated group. Using such a scheme there were 10 Worlds in the Non-Integrated category and 17 Worlds in the Moderately Integrated category. The same 9 Integrated Worlds were used here as were used in the first method.

Results

All 36 Worlds were scored according to the analytic system elaborated above, and a mean Integration score was obtained for each group. These means are as follows:

Non-Handicapped	22.91
Moderately Handicapped	14.20
Severely Handicapped	12.41

An analysis of variance yielded an F of 3.48 which is significant at the 5 per cent level of confidence. Individual t tests show that the difference between Non-Handicapped and Moderately Handicapped groups is significant at the 5 per cent level of confidence; the difference between Non-Handicapped and Severely Handicapped groups is significant at the 3 per cent level of confidence; and that the difference between the two handicapped groups is not significant. Thus one can conclude that children with a motor handicap have less integrative ability than non-handicapped children, and there is a tendency for increased severity of handicap to result in a corresponding decrease in integrative ability.

At this point an additional question was raised: Was this decrease due to the fact that the handicapped children merely made smaller but equally rich Worlds in terms of the complexity of the constituent units, or was there an impoverishment within the basic units of activity themselves? In order to obtain a measure of this, the proportion of units receiving high integration scores to the total number of scored units was calculated for each individual. Group means for each group were obtained and are as follows:

Non-Handicapped	.1469
Moderately Handicapped	.1012
Severely Handicapped	.0937

The standard error of the difference between proportions was obtained and used in calculating z scores between the groups. A one tail test of significance was used since the hypothesis predicted a decrease in integrative ability with motor handicap. When both handicapped groups were combined and compared with the Non-Handicapped group, a z of 1.67 was obtained which is significant at the 5 per cent level of confidence. A z score of 1.30 was obtained between the Non-Handicapped and Moderately Handicapped groups which is significant at the 10 per cent level of confidence, and a z score of 1.43 was obtained between the Non-Handicapped and Severely Handicapped groups which is significant at the 8 per cent level of confidence. This means that the handicapped children have fewer highly integrated units relative to the total number of units used in their Worlds, and there is a tendency for the proportion of highly integrated units to decrease with the increased severity of handicap. Thus one can conclude that the handicapped children have to rely on simple or moderately complex units in constructing their Worlds to a greater degree than the non-handicapped children.

Next the global ratings of the Worlds were analyzed. Following the first method, the number of children in the Non-Handicapped, Moderately Handicapped and Severely Handicapped groups falling into three categories of Integrated (I), Constricted (Cs) and Overinclusive (O) was tabulated. These results are presented in Table 1.

A chi square of 4.21 was obtained which is significant at about the 30 per cent level of confidence. When the Cs and O Worlds were combined in order to determine whether motor handicap was associated with an increase in any kind of deviate World, the results were even less significant. Thus there is no evidence of a relation between motor handicap and the production of deviate, in contrast to well integrated, Worlds.

TABLE 1

NUMBER OF NON-HANDICAPPED, MODERATELY HANDICAPPED AND SEVERELY HANDICAPPED CHILDREN FALLING INTO INTEGRATED, CONSTRICTED, AND OVERINCLUSIVE CATEGORIES

	Non-Handicapped	*Moderately Handicapped*	*Severely Handicapped*
Integrated	4	3	2
Constricted	4	4	8
Overinclusive	4	5	2

Finally, using the second method of grouping the global ratings of the Worlds, the number of children in the Non-Handicapped, Moderately Handicapped and Severely Handicapped groups falling into three categories of Integrated, Moderately Integrated and Non-Integrated Worlds was tabulated. These results are presented in Table 2.

A chi square of 2.90 was obtained which is significant at around the 25 per cent level of confidence. The results were also insignificant when the two handicapped groups were combined and compared with the Non-Handicapped group. Thus there is no evidence of a relation between motor handicap and the presence of extremely deviate Worlds.

TABLE 2

NUMBER OF NON-HANDICAPPED, MODERATELY HANDICAPPED AND SEVERELY HANDICAPPED CHILDREN FALLING INTO INTEGRATED, MODERATELY INTEGRATED, AND NON-INTEGRATED CATEGORIES

	Non-Handicapped	*Moderately Handicapped*	*Severely Handicapped*
Integrated	4	3	2
Moderately Integrated	6	4	7
Non-Integrated ...	2	5	3

DISCUSSION

In regard to the hypothesis tested, there is evidence that a motor handicap does lead to a decrease in integrative ability as measured by performance on the World test. Not only is there less ability in the absolute sense, but the proportion of highly integrated units to the total number of units used in constructing the Worlds is lower in handicapped than in non-handicapped children. In addition, there is a consistent tendency for increased severity of handicap to result in a corresponding decrease in integrative ability.

Since integrative ability has been defined in the present study in terms of richness and organization of ideation, one can conclude that the handicapped child can not bring the variety, flexibility and coherence of ideation to life situations that the non-handicapped child brings, and that he can not handle the complexities of life as effectively. Instead he prefers to take on fewer and simpler activities and to have a more limited area in which to function.

However, there are other important considerations. To begin with, the difference between handicapped and non-handicapped children was not as striking as one would expect on the basis of theoretical considerations.

When global ratings of the Worlds were used there were no significant differences between the groups in regard to either type or severity of deviation from good integration. At an empirical level this means one can not look at a World and say it is "typical" of the handicapped child. On a theoretical level it means that one can not conclude that non-handicapped children have good integrative ability and handicapped children have not, since in reality there is a great deal of overlap between the two groups. The soundest interpretation is that, at their best, the non-handicapped children can reach a level of complexity of organization in their dealing with problem situations which is above that of the handicapped children; and, at their worst, the handicapped children have to use meager or ineffectual solutions to a greater extent than do non-handicapped children. It also implies that the compensatory and restitutive mechanisms available to the handicapped child have been either underestimated or neglected. The present evidence indicates that they can maintain a constructive, integrative approach to the same extent as non-handicapped children even when they do not achieve the same level of complexity in their thinking; when they can not maintain this approach, they show the same kinds of disturbances as children with no motor handicap, and the handicap per se does not inevitably force them to become more constricted or impulsive than, say, an emotional disturbance in a non-handicapped child.

There are two possible explanations of these findings. As is known from other projective techniques the sacrifice of exploration and flexibility for the security of simple, easily mastered activities is often a consequence of anxiety. One can speculate that the decrease in integrative ability in the present study is due to the insecurity concomitant with a motor handicap and represents a defensive maneuver to protect the child from anxiety. The other alternative is that the simplification revealed in the study merely reflects the realistic simplification of a child's life resulting from the handicap. It is possible that such a child has to have more things done for him and thus does not have to use initiative as often as the non-handicapped child, and that his relations with parents and peers may be more limited and stereotyped, allowing for less free range of thought and activity. Because this simplified approach is perfectly adequate to the demands of his environment, he does not learn the complexities of adjustment that the non-handicapped child does.

Summary

In order to test the hypothesis that a motor handicap would result in a decrease in integrative ability, the World test was administered to a group of handicapped and non-handicapped children. Both an analytic scoring method and a global rating method were used to evaluate the complexity and organization of the Worlds. The results indicated that there is a sig-

nificant decrease in integrative ability in children with a motor handicap, with a tendency for increased severity of handicap to result in a corresponding decrease in integrative ability. Because of the large overlap between the groups, however, it was concluded that, at their best, non-handicapped children can reach a level of integration superior to that of the handicapped children, and, at their worst, handicapped children are forced to limit their activity more drastically than is necessary for non-handicapped children. There was no evidence that a motor handicap produces a particular kind of deviate thinking in the child, or necessarily forces him to extremes of deviate thinking.

References

1. Bolger, H., and Fischer, L. K. Personality projection in the World Test. *Amer. J. Orthopsychiat.*, 1947, 17, 117-128.

2. Wenar, C. The effects of a motor handicap on personality: I. The effects on level of aspiration. *Child Develpm.*, 1953, 24, 123-130.

SOCIO-ECONOMIC BACKGROUND OF CHILDREN WITH IMPAIRED HEARING [1]

MARTHA CRUMPTON HARDY
Elizabeth McCormick Memorial Fund

In connection with research on developmental problems of hard of hearing children, Fiedler (1) has reported that children with hearing defects were found to represent a segment of the population very low in socio-economic status. Her generalization that children so handicapped are "underprivileged in every sense of the word" has significant implications for workers in many disciplines with interests in the behavior and welfare of the child and warrants careful consideration of the supporting evidence. The purpose of this paper is to report results from further examination of her hypothesis in terms of the incidence of impaired hearing among school age children from widely differing socio-economic backgrounds.

The data consisted of records from the hearing and vision conservation program in the Chicago public schools covering the final 12 months (May 1950–June 1951) of a demonstration period during which key personnel were provided by private agencies to assist in launching the program. Procedures for case-finding of children with impaired hearing included group and individual pure-tone audiometric tests over six frequencies administered by technicians and examination in an otological screening clinic of cases failing to meet the minimum test standard. Children whose hearing threshold was below the 30 decibel level in two or more of the frequencies in either ear were referred to the screening clinic for diagnostic examination

[1] This study is a part of an exploratory investigation jointly sponsored by the Chicago Hearing Society, the Illinois Society for the Prevention of Blindness, and the Elizabeth McCormick Memorial Fund in cooperation with the Chicago Public Schools. Acknowledgment is made of the valuable assistance of the following who served as a Technical Committee: Otolaryngologists—Dr. George E. Shambaugh, Jr., Northwestern University Medical School; Dr. John R. Lindsay, University of Chicago, Department of Surgery; Ophthalmologist—Dr. J. Robert Fitzgerald, Loyola University, Stritch School of Medicine; Optometrist—Dr. Glenn H. Moore, Board of Northern Illinois College of Optometry; Pediatrician—Dr. Edward Lis, University of Illinois, College of Medicine; Medical Social Worker—Marie Waite, Illinois Eye and Ear Infirmary; Sociologist—Dr. Otis Duncan, University of Chicago, Community Inventory.

with the parent present. Those having a loss greater than 20 decibels but less than the standard for referral were designated "borderline" cases to be watched. The threshold, or zero decibel level, represents the point where hearing is 100 per cent present. For purposes of this study threshold acuity loss was classified as (1) borderline by test, (2) below minimum test standard, and (3) diagnosed as impaired hearing.

Subjects

The subjects included 36,692 pupils in the kindergarten, first, fourth, sixth, and eighth grades of 98 public elementary schools, of whom 1,014 or 2.8 per cent were referred for otological examination in the screening clinic. These were children with a suspected hearing loss, as indicated by individual audiometric tests. Differences between the grades in the number who failed to meet the minimum hearing requirements of the tests were small and well within the range of chance variations.

Incidence of hearing impairment as diagnosed by an otologist was estimated at 2.0 per cent of the pupil-population—2.0 per cent of kindergartners, 2.0 per cent of fourth graders, and 1.9 per cent of eighth graders. As many girls as boys were found to have impaired hearing. Although most of them had a history of earaches or discharging ears, in the large majority of cases their parents had not even suspected they could not hear normally until told of the hearing test results.

Two of every three of these pupils had a conductive type of impairment, which usually involved the middle ear. Both ears were affected about as frequently as one with normal hearing or only a mild degree of loss in the speech-hearing range in the better ear.

Of those referred for examination 99, or about 1 of every 10, had been under treatment for an ear disorder at some time prior to the hearing test at school.

Determination of Socio-Economic Background

Socio-economic background was estimated from neighborhood conditions as reflected by housing, extent of overcrowding, average monthly rental, median family income, and population character, according to the 1950 census reports for Chicago. Neighborhood was defined as a census tract. No single composite status index was formulated, rather each characteristic was used as a different aspect of the child's environment. Five classifications described variations from the over-all average for the city.

With respect to population, neighborhood classifications ranged from all white with at least 80 per cent native-born, to a high concentration of Negro —80 per cent or more. In 1950, 13.6 per cent of Chicago's population was Negro. Half of the neighborhoods represented in this study had an all white population and about a fifth of them were predominantly Negro.

Housing was categorized in terms of the extent of substandard living conditions in the area. Conditions varied from neighborhoods where there were few or no signs of deterioration of dwellings to blighted areas where at least half of the dwelling units were so deteriorated or so deficient in basic sanitation as to constitute a health hazard.

Some overcrowding was reported in all types of neighborhoods but in relatively few of the exclusively white sections was there more than is commonly found in Chicago. Where the population was mainly non-white, overcrowding was usual and living conditions in general were poor.

No attempt was made to secure a representative cross-section of Chicago's pupil-population. More than half of the children in this study were from the poorer sections of the city.

The design of the study called for determining socio-economic background from the home address but addresses were not readily accessible for pupils other than those who failed to meet minimum hearing or visual test requirements. Since a sample of 1,000 cases from 57 schools in widely differing types of neighborhoods showed close correspondence between conditions around these schools and those of the place of residence, socio-economic status of the pupils was determined by the characteristics of the school neighborhood, i.e., the census tract in which the school was located. For housing conditions this correspondence by the contingency coefficient was .86 for a 5×5 fold classification, as compared with the maximum value of .89. In terms of the favored neighborhoods, classifications were the same in 96 per cent and of the underprivileged in 97 per cent of the cases. Correspondence of other socio-economic conditions was of a similar degree.

Fiedler used social agency contacts of the families as one criterion for determining socio-economic status. Inquiry was made in the present study as to registration of families with community health and welfare agencies in the case of those whose children were referred for otological examination and of a sample representing children whose hearing was normal but who were identified by tests to have visual difficulties. This sample included 1,000 cases drawn from different grade levels and from various types of neighborhoods with distributions approximating those for hearing cases. The assumption here is that a history of contacts with social agencies reflects low socio-economic status as well as the existence of pressing personal problems.

Information on registrations with social agencies was obtained through clearance with the Social Service Exchange of the Welfare Council and concerned services provided by different agencies to any member of the family during the previous ten or twelve years, which covered the period since birth in the case of most of the children in the investigation. Agency services were grouped under six descriptive categories: financial assistance, health services, court services, child care and protective services, family services, and other types not included under these categories.

Findings

The general assumption in this investigation is that low socio-economic status with its concomitant health and social problems provides fertile ground for many of the infectious diseases whose after-effects may lead to ear disorders and hearing impairments. The hypothesis here tested is that children with impaired hearing are representative of a segment of the population very low in socio-economic status.

Chi-square tests were used to determine whether or not there was a significant relationship between hearing loss as shown by pure-tone audiometric tests or otological examination, and socio-economic status of the family as revealed by neighborhood living conditions or family history of social services. The 5 per cent level was used as the criterion of statistical significance.

Although there was a high degree of correspondence in the descriptions of a neighborhood by the several measures used to differentiate living conditions, relationships with hearing impairment were examined with respect to each aspect of the environment separately and with the population character of the area controlled, i.e., exclusively white areas or concentrated nonwhite. A further check on observed relationships was made in terms of a general rating where status categories of three of the four measures agreed in characterizing the neighborhood as (1) better than average, (2) average in Chicago, or (3) poorer. Fifty-six per cent of the neighborhoods represented were in one of these three categories and most of the children seen in the otological diagnostic clinic were from these neighborhoods.

The trend of the findings is indicated in Tables 1, 2 and 3, with chi-square values of significance shown in Table 4.

Housing Conditions

A large percentage of the neighborhoods represented are located in sections of the city where the worst housing conditions are concentrated and where overcrowding is most extensive. Reference to Table 1 will show that proportionately fewer cases of impaired hearing were found in neighborhoods where living conditions are substandard than in those where they are reported to be relatively good.

In the findings there is not the slightest suggestion of the greater frequency of impaired hearing, as revealed either by pure-tone audiometric tests or otological examination, among children from underprivileged neighborhoods, as judged by condition of dwelling units. Comparative differences were rarely large but with respect to comparisons concerned with children from exclusively white sections of the city the magnitude of the chi-squares indicates a significant negative relationship. Evidence from this study thus suggests the lower rate of hearing defects where living conditions tend to be unfavorable rather than the reverse.

TABLE 1

INCIDENCE OF IMPAIRED HEARING AMONG 36,692 PUBLIC SCHOOL PUPILS ACCORDING TO CHARACTER OF NEIGHBORHOOD—1950

Characteristics of the Neighborhoods	*Pupils Tested* (Number)	Incidence of Hearing Difficulties	
		Test Findings: Referred for Examination (Per Cent)	*Diagnosis: Impaired* (Per Cent)
Total	36,692	2.8	2.0
Housing Condition			
Better than Average	7,797	3.1	2.4
Average	8,243	2.9	2.1
Poorer than Average	20,652	2.6	1.8
Monthly Rent			
Above Average	3,569	3.8	2.8
Average	13,762	2.6	1.9
Less than Average	19,361	2.7	1.9
Population Character			
White	14,668	3.3	2.5
Mixed	10,103	2.5	1.9
Non-white	11,921	2.3	1.6
*Neighborhood Rating**			
Total	16,729	3.0	2.2
Above Average	2,736	3.8	3.0
Average	4,852	3.1	2.1
Poor	9,141	2.7	1.9

* Includes neighborhoods where characterizations by each of three or four aspects of living conditions were in agreement.

upper respiratory infections of various types which are said to be more likely to occur where overcrowding and poor sanitation are prevalent, the lack of definite evidence of the association of greater incidence of middle-ear

Analyses regarding the type of impairment showed a direct relation between housing conditions and impairment involving one or both ears. In the poorest neighborhoods a diagnosis of unilateral loss was more frequent and in the best neighborhoods bilateral loss, with differences significant at the 2 per cent level. In percentage terms, conductive impairment tended to be more frequently found where living conditions were substandard but such variations were not statistically significant—20 per cent level of confidence. Since conductive impairment is often the result of repeated

disorders with substandard housing conditions adds impressive emphasis to the absence of support for the hypothesis tested.

Monthly Rental and Income Level

Figures on rent refer to the average amount spent monthly for shelter without regard to the number of rooms or whether utilities and furnishings were provided. Income was defined as the net income received by families in 1949. According to the 1950 census reports, the median monthly rental in Chicago was $44.08 and the median family income $3,360.

Neither with respect to low monthly rentals or low income was there evidence of an association with frequency of impaired hearing among the pupil-population included in this study. Both high and low rentals were reported more frequently in sections where the population was exclusively white. Such differences in incidence of hearing impairments as were observed were suggestive of greater incidence where rentals were above average.

The trend of the findings as to income level likewise indicated greater frequency of impaired hearing among children from families in the higher income brackets with a progressively decreasing proportion in the lower income groups. In neighborhoods where the median income was in excess of $4,500, the percentage with hearing below minimum requirements of the test was 3.6; where the median was less than $2,000 this percentage was 2.1.

Relationships with low economic status were further examined in terms of public assistance rates in the different community areas in Chicago. The city is divided into 75 community areas comprising 935 census tracts, here designated neighborhoods. Assistance rates were available by specific age groups and place of residence for persons receiving cash grants. The 18 areas with the highest rates in 1952 for the population under 18 years of age contained 26 per cent of this age group, but 86 per cent of the public assistance recipients of this age group (2). One of every four children included in the present study was living in a high assistance rate area. Incidence of hearing impairment was found to be lower in such areas than where few or no families with children under 18 years of age were receiving public assistance—2.1 per cent where the public assistance rate was high as compared with 3.7 per cent where this rate was low.

Population Character of Neighborhood

The population character of the different neighborhoods was described as an all white area, mixed population area, or concentrated non-white area. White and mixed classifications were divided into subgroups to show nativity in the white areas and the proportion of non-white in the mixed areas, i.e., less than average in Chicago or average but not predominantly non-white.

Analyses of the screening results reveal a direct association between incidence of impaired hearing and the population character of the neighborhood, with the highest frequency of impairment found among native-born white and the lowest among children from concentrated non-white areas. The fact that most of the white children were from sections of the city where living conditions are relatively good and the large majority of the non-white were from underprivileged sections make these findings peculiarly pertinent to the problem investigated.

TABLE 2

COMPARISONS OF FREQUENCIES OF CERTAIN TYPES OF HEARING IMPAIRMENTS AMONG SCHOOL CHILDREN FROM DIFFERENT NEIGHBORHOOD CONDITIONS

		Percentage Having Specified Diagnosis					
Characteristics of the Neighborhoods	*Pupils with Impaired Hearing*	*Loss in One Ear*	*Loss in Both Ears*	*Conductive Impairment* Total	*Middle Ear Involved*	*Perceptive Impairment Inner Ear Involved*	*High Tone Loss*
Total	740	52.7	47.3	66.3	52.7	18.8	14.9
Housing Condition							
Better than Average	190	46.0	54.0	58.1	48.4	15.0	26.9
Average	169	52.1	47.9	61.6	47.6	20.7	17.7
Poorer than Average	381	56.2	43.8	72.6	57.2	19.8	7.6
Monthly Rent							
Above Average	98	51.5	48.5	71.3	59.6	9.6	19.1
Average	267	50.4	49.6	60.1	46.9	20.9	19.0
Less than Average	375	54.6	45.4	69.5	55.0	19.6	10.9
Population Character							
White	362	48.2	51.8	59.6	49.1	18.5	21.9
Mixed	188	54.9	45.1	76.2	59.7	13.8	10.0
Non-white	190	58.9	41.1	69.3	52.7	24.2	6.5
*Neighborhood Rating**							
Total	362	51.9	48.1	65.4	50.4	18.4	16.2
Above Average	83	48.8	51.2	60.5	48.2	11.1	28.4
Average	103	45.6	54.4	61.4	49.5	16.8	21.8
Poor	176	57.1	42.9	70.2	52.1	22.8	7.0

* Includes neighborhoods where characterizations by each of three or four aspects of living conditions were in agreement.

The trend of the findings was consistent regardless of whether they concerned borderline caces (by tests), those failing to meet the minimum hearing requirement with referral for otological examination, or those with a diagnosis of impaired hearing. Statistically the evidence is significant in showing more cases of impaired hearing among white children than would be theoretically expected and fewer among the non-white. Inspection of the mid-section of Table 1 will indicate the marked consistency of the differences between classifications as to population character of the neighborhood. While the categories may not correctly describe an individual child's race, the population of neighborhoods designated as non-white areas was almost exclusively non-white with 80 per cent to 99 per cent Negro, and that of the white areas represented was entirely white.

Differences were also indicated regarding types of hearing impairment and history of ear trouble. In the white areas bilateral loss was more frequent and in the non-white areas unilateral loss. Findings in percentage terms based on records in which such diagnoses were noted are shown in the mid-section of Table 2. Here again the differences were statistically significant. Histories of earaches and discharging ears likewise were reported more frequently for white children than for those from non-white areas. Since facts concerning history of ear disorders were elicited by direct questioning at the time of the clinic visit the information is probably as accurate as any other type of health history so obtained, and no less in the case of the non-white than the white mother. It is noteworthy in this connection that about as many of the non-white as white who were seen in the otological screening clinic had been under medical care for an ear disorder at some time prior to their examination at school—13.0 per cent of the white, and 12.4 per cent of the non-white for whom such information was available.

If it can be assumed that in the Chicago area the population character of a neighborhood tends to reflect the socio-economic status of its residents, with exclusively white neighborhoods where most of the families are so-called Old Americans representing the highest status category and concentrated non-white neighborhoods the lowest, then these findings clearly fail to support the generalization that the majority of children whose hearing is impaired are from families of very low socio-economic status.

General Character of Neighborhood

In terms of a composite descriptive rating based on the extent of dilapidation, overcrowding, amount of monthly rental, and/or family income level when the population character was controlled, the analyses again yielded no indication of an association between substandard neighborhood conditions and impaired hearing among school children living under such conditions. Since there was a high degree of correspondence between the dif-

ferent measures of socio-economic status, the trend of these results was to be expected from findings previously cited. The results clearly warrant the conclusion that among elementary school pupils from widely differing living conditions in a metropolitan community hearing impairment, as determined by pure-tone audiometric tests and confirmed by otological examination, was observed with no greater frequency among children from underprivileged than from favored neighborhoods.

Social Agency Contacts

Family background information was obtained on 1,865 individual children regarding history of registration of some member of their immediate families with community social agencies and the number and type of agencies with whom they have had contact. Included among these cases were 605 children who had a diagnosis of impaired hearing, 105 who had been referred to the otological screening clinic but upon examination were found to have "no loss" of hearing, and 1,155 with visual difficulties whose hearing was normal. More than half of the vision cases were living in sections of the city classified as average or better. A history of social agency services was reported for two of every five of the families for whom information was secured—36 per cent of the impaired hearing cases, 36 per cent of the "no loss," and 42 per cent of the visual defect cases. These records reveal such histories were no more frequent among children with hearing defects than among those with normal hearing.

In the majority of cases the contact had been with only one agency although 6 per cent of the families had been registered with at least four different agencies for some type of service during the years 1940-52 inclusive. The three types of services reported most frequently were for financial aid, health services, and court services. Histories of financial assistance were observed as often in the case of children with normal hearing as of those with impaired and were most frequent among the non-white. Among the impaired hearing cases from non-white neighborhoods 31 per cent had a history of financial assistance from welfare agencies; among the normal hearing with visual difficulties 54 per cent. These were the only comparisons with respect to social agency contacts where differences were statistically significant.

Table 3 shows the types of social agency contacts reported for families of children whose hearing was below test requirements compared with those with normal hearing who failed to meet the vision test standards. Findings are shown in percentage terms for the total cases cleared and for those living in exclusively white neighborhoods and in neighborhoods predominantly non-white. While all of these children had been identified as having either hearing or visual difficulties, the lower status of those with hearing loss is not indicated in these findings.

TABLE 3

COMPARISONS OF SOCIAL AGENCY CONTACTS OF 1,865 FAMILIES OF IMPAIRED HEARING OR VISION CASES

History of Agency Services	HEARING OR VISION BELOW TEST REQUIREMENTS					
	Total Sample*		White Neighborhoods		Non-White Neighborhoods	
	Hearing Cases	*Vision Cases*	*Hearing Cases*	*Vision Cases*	*Hearing Cases*	*Vision Cases*
Number in Sample	710	1,155	375	599	163	279
No Agency Contacts	451	673	273	465	83	76
Agency History	259	482	102	134	80	203
Percentage with History†	36.5	41.7	27.2	22.4	49.1	72.8
Financial Services	16.3	24.4	7.7	6.8	31.3	54.1
Health Services	15.5	11.8	12.5	9.8	14.1	14.7
Court Services	14.5	11.7	10.7	8.5	19.6	14.3
Child Care Services	9.2	5.8	8.3	4.0	7.4	7.9
Family Services	8.6	5.2	6.7	5.5	12.3	4.7
Other	3.5	2.9	3.7	2.2	3.7	2.9

* Includes 172 Hearing cases and 277 Vision cases from mixed population neighborhoods.

† Percentages based on size of total sample.

DISCUSSION

Socio-economic background of children with impaired hearing was examined in terms of certain characteristics of their physical environments and the need of their individual families for special community services as revealed by registration with social agencies. While the place of residence may not afford an adequate index to a family's status, neighborhood conditions indicate the "climate" in which these children were growing up and substandard living conditions assuredly tend to militate against healthy development. The length of time the families had been living under environmental conditions of the character of their 1950 addresses was not ascertained but from available information it would appear that the population of the poorest neighborhoods in Chicago tends to be about as stable as that of favored sections. According to census reports, in the exclusively white areas represented in this study 87 per cent of the families in neighborhoods here characterized as underprivileged had not changed their place of residence during 1949 and 1950 as compared with 88 per cent in the best sections.

It is a logical assumption that if the effects of substandard living conditions directly or indirectly bear a close relation to impaired hearing in childhood, they should be reflected in comparative results of incidence in widely differing neighborhoods. The findings in this study show neither in consistent trend nor in statistically significant differences evidence which would suggest higher incidence of impaired hearing where living conditions were poor. Even where dwelling units were so deteriorated or so deficient in basic sanitation as to constitute a health hazard and the entire neighborhood was characterized as blighted, the frequency of hearing impairment among these children of school age was no greater, if as great, as among children living in the best residential sections. Children with impaired hearing were discovered in all types of neighborhoods and, in metropolitan Chicago, cannot be said to represent any special segment of the population.

Home visits to 50 cases so selected as to give a cross-section of the different socio-economic categories confirmed the classification of living conditions in three of every four of the cases visited. Generally such variations as occurred were in areas where signs of deterioration were less extensive, and where in the main poorer conditions were the result of subdivisions of large apartments into small dwelling units with the consequent overcrowding. Individual home conditions were considered as better than those characterizing the neighborhood in eight of these cases and poorer in six.

Detailed examination of evidence regarding low socio-economic status of families of children with impaired hearing as revealed by their contacts with social agencies likewise yielded no indication of unique family background in this respect. While comparable data from the community as a whole were not available, in the light of findings from a group with normal hearing who were identified as in need of eye care, the majority of whom were from average or better types of neighborhoods, it seems reasonable to assume that the extent of agency contacts of the impaired hearing cases differed little from what might be expected in this metropolitan community during the 10 year period covered. Apparently whatever situations these histories of contacts with health and welfare agencies may reflect their impact had been felt by the normally hearing child as well as by the child with impaired hearing.

The level of confidence of the relation of various social and economic aspects of family background to incidence of hearing impairment among approximately 37,000 Chicago school children as determined by pure-tone audiometric tests or otological examination is shown in Table 4. In each correlation the direction of the results pointed to the higher status of the impaired hearing cases and the greater likelihood of loss in both ears.

All hearing tests were administered by a staff of technicians who were sent from a central office to the individual schools. During the period covered by this study the selection of schools where tests were to be given was determined by requests from the principals. Machines were calibrated

TABLE 4

SIGNIFICANCE OF VARIATIONS IN INCIDENCE OF IMPAIRED HEARING WITH MEASURES OF SOCIO-ECONOMIC STATUS

Hearing Loss	*Aspects of Socio-Economic Status*	*Chi-Square Value*	*Degree of Freedom*	*p*
Acuity below test standard	White—Non-white neighborhoods	24.14	1	.01
Diagnosis of loss	White—Non-white neighborhoods	4.74	1	.03
Bilateral—Unilateral	White—Non-white neighborhoods	5.76	1	.02
Acuity below test standard	Housing	6.97	3	.10
Acuity below test standard	Housing (White neighborhoods)	8.74	3	.05
Diagnosis of loss	Housing	9.75	3	.05
Bilateral—Unilateral	Housing (White neighborhoods)	10.66	3	.02
Hearing—Vision defect	Financial—Other agency services	17.28	2	.01
Hearing—Vision defect	Financial—Other agency services (Non-white cases)	26.77	2	.01

and headphones checked before each day's testing was begun and the equipment spotchecked during the day. Special effort was made to keep the testing and otological examination procedures comparable throughout all schools. It is therefore most unlikely that the comparative findings between good and poor neighborhoods were unduly influenced by variations in standardization of procedure, interest of teachers, or cooperation of the pupils.

Only a relatively small proportion of the families of these impaired hearing cases had even suspected the child was having difficulty hearing until so advised by the school, although the loss was bilateral about as often as unilateral and in some cases was severe. Hearing is an important part of the dynamics of behavior and prevention of impairment consists not only of the reversal or alleviation of clinical symptoms of the ear, but also of the mitigation of collateral symptoms related to behavior and development. While the impaired hearing cases in this study were no less likely to be from favored as from underprivileged neighborhoods, what was done about their impairment varied directly with the character of their living conditions but appeared to be unrelated to previous knowledge of the hearing difficulty. Fewer children from the underprivileged neighborhoods were given an opportunity for needed medical care than those from the better sections of the community. Children from non-white sections were as likely to be taken for the advised care as those from exclusively white sections when their living conditions were similar.

The results on follow-up care are in agreement with those reported in the literature on medical care and remedial services in showing a direct relation

between care obtained and socio-economic status. The surprising finding in this study is that so large a number of children with impaired hearing from so-called middle class families who were living in the better residential sections where incomes were considerably above average failed to receive the recommended medical care. According to the available records less than half of the children from the higher status levels were taken for care, and not all of these were seen by an otologist as their condition required. These findings offer convincing evidence of the extensive need among families of all economic classes for intelligent understanding of the importance of hearing in the social adjustments of everyday living and of what impaired hearing means.

Reactions of parents to a diagnosis of impaired hearing will vary with individuals but will undoubtedly influence what is done about the child's problem. From interviews with parents who had failed to follow through on the care advised by the examining otologist it was apparent that numbers of them considered impaired hearing as something to be tolerated with little acceptance of the possibility of remedial or rehabilitative steps to offset the communicative disorder; others refused to accept the diagnosis and expressed no concern for further evaluation of the condition.

Research at both basic and applied levels on aspects of behavior and development of children with any degree of hearing loss needs to be carried on as well as in areas of prevention and remedial measures. With the great advances in medical treatment, the rapidly developing field of audiology and in modern concepts of the dynamics of individual development, the horizon of what can and should be done for children with impaired hearing has been tremendously widened. The major problems center around early detection of the child with ear disease, accurate medical diagnosis of his impairment with treatment as soon as possible, and adequate provisions for special guidance and services to help him realize his potentialities. Implications of the Chicago findings suggest it is with respect to these problems that relationships with socio-economic status can be identified, but that such relationships appear to be of less importance than parental attitudes toward hearing impairment and the family's value systems.

Summary

The hypothesis that children with impaired hearing are representative of a segment of the population very low in socio-economic status has been examined from the point of view of incidence of hearing loss among children from sharply contrasting socio-economic backgrounds. The results fail to establish the hypothesis, whether socio-economic status was determined by characteristics of the neighborhood environment or aspects of living conditions of individual families and their need for community services as revealed by contacts with social agencies. Where differences were

observed they suggested a lower incidence of hearing defects among children from underprivileged neighborhoods than from the favored.

References

1. Fiedler, Miriam F. A study of the socio-economic status of a group of public school children with hearing defects. *Child Develpm.*, 1951, 22, 193-198.

2. *Public assistance rates for the 75 community areas and the city of Chicago, October 1952.* Research Department, Welfare Council of Metropolitan Chicago. *Statistics,* 1953, 20, No. 11.


C51